1970

An investigation of the robustness of Student's t-test using Beta and Erland distributions.

Nomani, M. Rizwan

Monterey, California ; Naval Postgraduate School

AN INVESTIGATION OF THE ROBUSTNESS OF "STUDENT'S" T-TEST USING BETA AND ERLANG DISTRIBUTIONS

M. Rizwan Nomani

United States
Naval Postgraduate School

THESIS

AN INVESTIGATION OF THE ROBUSTNESS
OF "STUDENT'S" t-TEST USING
BETA AND ERLANG DISTRIBUTIONS

by

M. Rizwan Nomani

September 1970

An Investigation of the Robustness
of "Student's" t-test Using
Beta and Erlang Distributions

by

M. Rizwan Nomani
Major, United States Army

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 1970

ABSTRACT

The robustness of "Student's" t-test with regard to the assumption of normality is investigated. This is accomplished by empirically developing the distribution of a new statistic from a large number of samples of varying sizes from the Beta and the Gamma distributions using different values of the parameters. The various significance levels of this new distribution are then compared with the corresponding significance levels from "Student's" t distribution using an IBM 360 computer. A comparison of the distribution frequencies with the standard normal distribution frequencies is also presented.

TABLE OF CONTENTS

I. INTRODUCTION

In practical work, an Operations Research Specialist is generally involved with the problems of bringing the theoretical structures into some degree of correspondence with the situations of practical experience. One of the common methods of achieving this objective is by hypothesis testing. The hypothesis testing method makes use of a variety of statistical techniques and procedures involving both parametric and non-parametric tests. In the derivation of most of the parametric tests, it is usual to assume a form of mathematical model involving some specific probability distribution and then to select some statistical criterion that is sensitive to change in the specific factors tested. This type of criterion is generally known as the "power of the test".

However, another desirable requirement of any statistical test is that it be insensitive to change in the underlying assumptions. This property of the test is generally referred to as the "robustness" of the test. Common questions probing into the robustness of a test are: When the actual distribution is not known, which statistical tests can be used with less hesitation?; and What percentage of error would be incurred if the underlying assumptions are violated? Past research and subsequent literature indicate that the property of robustness is generally not satisfied by the parametric tests. This is substantiated by research conducted by G.E.P. Box [7] and R. C. Geary [8].

One of the most common parametric tests used for hypothesis testing is "Student's" t-test. This test was developed by

William Sealy Gosset in 1908 when he was working for Guinness Brewery in Dublin, Ireland. The discovery of the t-test was a direct result of the peculiar problems of brewing with its variable material and its susceptibility to temperature changes. A number of experiments conducted at the brewery emphasized the limitations of large sample theory and revealed the necessity for a correct method of treating small samples. It was, therefore, the circumstances of Gosset's work which led to his discovery of the Theory of Errors and the distribution of the Sample Standard deviation, which was developed later into the well known t-test.

II. "STUDENT'S" t DISTRIBUTION

As was mentioned earlier, the Theory of Errors finds its origin in the fact that the accuracy of the mean of a number of observations may be estimated from the discrepancies observed among the individual values used in obtaining the mean. When the samples are drawn from a normal population, having a variance (mean squared deviation) equal to σ^2, it can be shown that the mean of N observations $\bar{x}$ is also normally distributed with variance $\sigma^2_{\bar{x}} = \frac{\sigma^2}{N}$. Thus if the variance of the mean is known, then the desired information regarding the distribution of the mean can be easily determined. In order to test a hypothesis that the population has mean μ, one would merely need to calculate $t = \frac{\bar{x}-\mu}{\sigma/\sqrt{N}}$ and then the integral

$$I = \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{\frac{-t^2}{2}}\, dt$$

would give the probability that a more discrepant value would occur. If the value of I so calculated turns out to be a small quantity, such as .01, one should conclude with some confidence that the hypothesis was not, in fact, true of the population sampled.

In a majority of the cases in which such tests are required, one can find no prior knowledge of the variance of the population. The variance of the population can be estimated, however, from the sample itself by the following relationships: If $X_1, X_2, \ldots, X_N$ represents a sample; then $s^2 = \frac{1}{N-1} \sum_{i=1}^{N} (X_i - \bar{x})$ is an estimate of the unknown variance σ^2, and $\bar{x} = \frac{1}{N} \sum_{i=1}^{N} X_i$ is the sample mean and an estimate of the population mean.

W. S. Gosset, in his fundamental paper of 1908, showed that even though s as calculated above is a good estimate of σ, it would be erroneous to assume that the statistic $t = \frac{\bar{x}-\mu}{s/\sqrt{N}}$ would be normally distributed or that the significance of an observation could be accurately tested by using the normal probability theory. Assuming the independence of s^2 and $\bar{x}$, Gosset was able to develop the exact distribution of this statistic, which in its modern form is known as "Student's" t distribution. It states that if X and U are two independently distributed random variables such that X is normally distributed with mean μ and variance σ^2, and U has chi-square distribution with N degrees of freedom, then the ratio $t = \frac{X-\mu}{\sigma} / \sqrt{U/N}$ has the distribution of "Student's" t, and its density function is given by

$$f(t) = \frac{\Gamma(\frac{N+1}{2})}{\sqrt{\pi N}\,\Gamma(\frac{N}{2})} \cdot \frac{1}{[1 + \frac{t^2}{N}]^{\frac{N+1}{2}}}, \quad \text{where } -\infty<t<\infty$$

From this definition, it can be seen that "Student's" distribution depends on N only and is independent of both parameters of the normal distribution sampled. It is further noted that the derivation of "Student's" t distribution depends on the fact that the random variable X be drawn from a normal population.

One of the common uses of "Student's" t-test is to compare the mean of a population with some standard. In practical problems, however, it is often the case that the distribution

of the population being sampled is not known or is not normal. It is therefore desirable that some measure of the degree of error be obtained for "Student's" t-test when the assumption of normality of the sample is violated.

Literature reveals that a great deal of research has been directed towards many aspects of "Student's" distribution. In 1925, R. A. Fisher very stringently verified the independence of the sample mean $\bar{x}$ and the sample variance s^2. This was one of the basic assumptions in the development of "Student's" distribution [2]. In 1936, R. C. Geary investigated the distribution of "Student's" t ratio for the samples drawn from a slightly asymmetrical universe by developing asymtotic formulae for the moments of the t distribution. He showed that for symmetrical, non-normal populations, "Student's" distribution gives more accurate results than skewed sampling distributions. (4) In 1948, A. K. Gayen, using the moments techniques, studied the behavior of "Student's" t and obtained results similar to those obtained by R. C. Geary in 1936 [5]. In more recent work, Cucconi 0 developed a simple relation between the critical values of "Student's" t and the degrees of freedom. Through this relation, in addition to deriving the critical values of t for any number of degrees of freedom, he showed that it is possible to substitute "Student's" criteria in the verification of statistical hypothesis more easily than using a standardized normal deviation [9]. These are but a few examples of the diversified attention given to "Student's"

t-test. The aim of this paper is to investigate the robustness of "Student's" t-test when the sampling data is drawn from the Beta and the Gamma distributions.

III. METHODS AND PROCEDURES

In order to investigate the robustness of t-test, it was decided to select two well known distributions with a wide scope of application in the real world. The two distributions selected are the Gamma and Beta distributions. A brief description of these distributions follows.

A. GAMMA DISTRIBUTION

A continuous random variable Y is said to have the Gamma distribution if its density is

$$f(y) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} y^{\alpha-1} e^{\frac{-y}{\beta}} \quad , \quad 0<y<\infty$$

$$= 0 \quad , \quad \text{otherwise}$$

It should be noted that this is a two parameter family of distributions with parameters α and β. The distribution is only defined when the value of $\alpha>-1$, and the value of $\beta>0$. The mean and variance of this distribution are respectively $\alpha\beta$ and $\alpha\beta^2$. The Gamma distribution is frequently applied to the solution of the queuing and inventory control problems. A graphical representation of the Gamma distribution for various values of its parameters is given in Appendix A.

One of the special and most often used family of Gamma distributions is called the Erlang distribution. In order to obtain this distribution, the values of the parameter α in the Gamma distribution are restricted to positive integer values and this positive integer is denoted by K. If one replaces

β by $\frac{1}{\lambda}$, then the distribution becomes

$$f(y) = \frac{\lambda(\lambda y)^{K-1} e^{-\lambda y}}{(K-1)!} \qquad \text{for } y \geq 0$$

$$= 0 \text{ , otherwise}$$

It is possible to show that the latter f(y) can be generated by the sum of K variates Z having the exponential density

$$f(z) = \lambda e^{-\lambda z} \qquad \text{when } z > 0$$

$$= 0 \text{ , otherwise}$$

It should be noted that when K = 1, we have the exponential density as a special case. The mean, variance and mode of the Erlang distribution are $\frac{K}{\lambda}$, $\frac{K}{\lambda^2}$ and $\frac{K-1}{\lambda}$ respectively. The great usefulness of this distribution stems from the fact that it is a large family of distributions permitting only non-negative values. The waiting time and the service time distributions in the queuing theory can, therefore, be reasonably approximated by an Erlang distribution.

B. THE BETA DISTRIBUTION

A continuous random variable Z is said to have a Beta distribution when

$$f(z) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} z^{\alpha-1}(1-z)^{\beta-1} \quad \text{when } 0 \leq z \leq 1$$

$$= 0 \text{ , otherwise}$$

This density is a function of the two parameters α and β, both of which are positive constants. The mean and variance of

this distribution are given by $\frac{\alpha}{\alpha+\beta}$ and $\frac{(\alpha+1)\ (\beta+1)}{(\alpha+\beta+2)^2(\alpha+\beta+3)}$. A graphical representation of the Beta distribution for various values of the parameters in furnished in Appendix B.

It may be pointed out that when $\alpha=\beta=0$ the distribution becomes the uniform distribution over the unit interval. The Beta distribution, though in general defined over the unit interval, can also be defined over any interval (a, b) as follows

$$f(z) = \xi(z-a)^{\alpha}(b-z)^{\beta} \qquad \text{where } a \leq z \leq b$$
$$= 0 \text{ , otherwise}$$

where $\xi = \int_a^b (z-a)^{\alpha}(b-z)^{\beta}dz$

Primarily because the Beta distribution is defined on a finite interval, it is widely used in the study of critical path scheduling techniques such as PERT. It may be recalled that PERT is a method of reflecting the uncertainties associated with development-type tasks, and its use generally results in providing a more realistic outlook for task accomplishment than the conventional systems provided in the past. In order to obtain the critical path through the network, it is necessary that a distribution defined between the pessimistic and optimistic time estimates adequately represent the distribution of time required to perform an activity. Since the Beta distribution restricts the probability distribution to a finite range for varying values of the parameters, it is widely used in the critical path scheduling.

C. GENERATION OF DISTRIBUTIONS

The inverse probability integral transformation technique is used to produce random variates having the exponential distribution. According to this technique, the cumulative distribution function for the uniform probability density function on the interval (0,1) is equated to the cumulative distribution function for the desired probability density function in order to obtain the necessary transformation. Thus if RN denotes a random variable on the interval (0,1), and λ is the desired parameter of the exponential distribution; then the random variable obtained from the transformation $X = \frac{-\ln(RN)}{\lambda}$ is exponentially distributed with a mean $\mu = \frac{1}{\lambda}$.

1. Erlang Distribution

In order to generate the Erlang distribution, an extension of the above procedure is utilized by first generating exponential random variables with mean $\frac{1}{\lambda}$. Then K of these exponential variates are summed using the relationship

$X = \sum_{i=1}^{N} \frac{\ln(RN)i}{-\lambda}$ to produce the desired Erlang distribution with mean $= \frac{K}{\lambda}$ and variance $\frac{K}{\lambda^2}$.

2. Beta Distribution

The Beta distribution is generated using the following theoretical ideas. Consider two random samples $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_m$ such that the random variables X and Y are identically distributed exponential with parameter λ. If we

form the statistics $Z_1=\sum_{i=1}^{n} X_i$ and $Z_2=\sum_{i=1}^{m} Y_i$ then

$$Z_1 \sim \Gamma(z_1;\ n,\mu)$$

$$Z_2 \sim \Gamma(z_2;\ m,\mu)$$

and their joint density function is given by

$$g(z_1,z_2) = \frac{z_1^{n-1} z_2^{m-1} e^{-(\frac{z_1+z_2}{\mu})}}{\Gamma(n)\Gamma(m)\mu^{m+n}}, \quad 0<z_1<\infty \text{ and } 0<z_2<\infty$$

Now using the transformation $W_1 = \frac{Z_1}{Z_1+Z_2}$ and $W_2 = Z_1+Z_2$, the joint density of W_1 and W_2 is given by:

$$h(w_1,w_2) = \frac{(w_1w_2)^{n-1}w_2^{m-1}(1-w_1)^{m-1}}{\Gamma(n)\Gamma(m)\mu^{m+n}} e^{-\frac{w_2}{\mu}} \cdot w_2$$

Integrating over the domain of W_2, the marginal density of W_1;

$$h(w_1) = \int_0^\infty h(w_1,w_2)\ dw_2$$

i.e., $$h(w_1) = \frac{w_1^{n-1}(1-w_1)^{m-1}}{\Gamma(n)\Gamma(m)\mu^{n+m}} \int_0^\infty w_2^{n+m-1} e^{-\frac{w_2}{\mu}}\ dW_2$$

Using the substitution $U = W_2/\mu$ and $dW_2 = \mu\ du$ we get:

$$h(w_1) = \frac{w_1^{n-1}(1-w_1)^{m-1}}{\Gamma(n)\Gamma(m)} \int_0^\infty u^{n+m-1}e^{-u}du$$

$$= \frac{\Gamma(n+m)}{\Gamma(n)\Gamma(m)} w_1^{n-1}(1-w_1)^{m-1}$$

We observe that $W_1 = \frac{Y_1}{Y_1+Y_2}$ has the Beta distribution with $\alpha=n$, $\beta=m$, and this Beta distribution is independent of the parameter of the exponential distribution provided that the two Gamma distributions are generated by random numbers having identical exponential distributions. Thus to generate the Beta distribution, two Gamma distributions with parameters $(K_1,\frac{1}{\lambda})$ and $(K_2,\frac{1}{\lambda})$ are generated using the above techniques. From these distributions, the ratio $W_1 = \frac{Y_1}{Y_1+Y_2}$ is formed to obtain a Beta distributed random variable with mean $= \frac{K_1}{K_1+K_2}$ and variance $\frac{K_1\ K_2}{(K_1+K_2)^2(K_1+K_2+1)}$.

D. SIMULATION TECHNIQUES

The accuracy of generating random observations from a probability distribution depends upon the characteristics of the uniform random number generator. At least the superficial characteristics of the random number generator are ascertained by the use of chi-square test for goodness-of-fit at 95% and 99% significance levels. For the chi-square test, the unit interval is divided into fifty equal intervals. Each generated random number is assigned to one of the fifty categories according to its size. From this data, a measure of the discrepancy existing between observed and expected frequencies is obtained using the relationship $X^2 = \sum_{i=1}^{T} \frac{(o_i-e_i)^2}{e_i}$ where o_i denotes the observed frequency of the ith category, and e_i

denotes the expected or theoretical frequencies of this occurence; T denotes the total number of categories. The number of degrees of freedom for the statistics is equal to T-1, where T is the number of categories or classes.

For the purpose of this project, four simulated distributions were obtained for each type of distribution sampled. Each simulation utilizes 5000 samples and a fixed sample size. In the discussion throughout this paper, sample size is referred to as the number of data points within each sample used to calculate the t statistic and the number of samples denotes the number of replications for fixed values of the parameters within each cycle. In order to obtain different shapes of the distribution, five different sets of values of the parameters are then used within each simulation.

Since the basic techniques and procedures employed in each of the four simulations for the Erlang and the Beta distributions are essentially the same, it is considered adequate to discuss the methodology of one simulation for each distribution in detail.

1. Erlang Distribution

Simulation 1 for the Erlang distribution consists of four steps. Step I consists of five cycles. For each cycle, 5000 samples of size 5 are independently generated from the Erlang distribution using the values of the parameters as shown in Table 1 below. For the purpose of discussion, these sampling distributions are referred to as $\gamma(K,\lambda)$

TABLE 1

Cycle #	Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5
Value of Parameters $\gamma(K,\lambda)$	$K = 4$ $\lambda = 3$	$K = 5$ $\lambda = 3$	$K = 6$ $\lambda = 3$	$K = 7$ $\lambda = 3$	$K = 8$ $\lambda = 3$
Mean	1.333	1.667	2.0	2.333	2.667
Standard Deviation	.6669	.7408	.8200	.8624	.9368

To illustrate the use of Table 1, 5000 samples of size 5 are generated during cycle 1 using parameters K =4, λ =3. The mean and standard deviation of this Erlang distribution are given in rows 3 and 4 of Table 1. During Cycle 2, another 5000 samples of size 5 are independently generated from the Erlang distribution using parameters K =5, λ =3. The mean and standard deviation of this distribution are 1.667 and .7408 respectively. These values are furnished in rows 3 and 4 of Table 1 under the heading Cycle 2. The same process is repeated for Cycles 3, 4, and 5 using the appropriate parameters.

In Step II, for each sample the statistic $t = \frac{\bar{x}-\mu}{s/\sqrt{N}}$ is calculated where $\bar{x}$ is sample mean, s is sample standard deviation, N = sample size (5), and μ is population mean (1.33). The absolute values of this statistic are then compared with the corresponding values from the "Student's" t-distribution

with N-1 degrees of freedom and at significance levels $\alpha=.8$, .5, .2, .1, .05, .02, .01, and .001. These values of "Student's" t-distribution with varying degrees of freedom and significance levels are generally referred to as the critical values t_α of t-distribution and are the solutions to the equation $1-\frac{\alpha}{2}= \int_{-\infty}^{t\alpha} f(t)dt$ for various values of alpha. The number of t statistics whose absolute values are greater than the corresponding critical values at each significance level are recorded for comparison purposes.

In order to fully understand the comparisons between the actual significance levels and the significance levels obtained from the Erlang distribution, it is necessary that some knowledge regarding the comparisons between the distribution frequencies be known. This is accomplished in Step III by transforming the total number of data points for the Erlang distribution in each cycle (5000 X 5 = 25,000) using the transformation $\frac{X_i-\mu}{\sigma}$, $i = 1,25000$, where μ and σ are the sampling distribution mean and standard deviation respectively. The transformed distribution is called $\gamma(0,1)$. It may be pointed out that if the sampling distribution were normal, with mean μ and standard deviation σ then this transformation $\frac{X-\mu}{\sigma}$ would reduce the sampling distribution to what is called Standard Normal, mean zero, variance unity. This is generally denoted by N(0,1). The transformed distribution, $\gamma(0,1)$ is then divided into four intervals of width unity on either side of assumed mean of zero, and the frequency counts within each

interval are recorded. These counts are then compared with the expected number of frequency counts within each interval assuming that the sampling distribution was normal.

In Step IV, the chi-square goodness-of-fit test, using a significance level of 0.05, is conducted on the results of the observed significance levels and the distribution frequencies as calculated in Steps II and III above. This concludes Cycle 1 of Simulation 1. For Cycles 2, 3, 4, and 5, the entire procedure as outlined above is repeated by changing the parameters of the $\gamma(K_1,\lambda)$ distribution in accordance with Table 1.

Simulations 2, 3, and 4 for the Erlang distribution use exactly the same techniques as outlined under Simulation 1 except that the sample sizes are 10, 15, and 20 respectively.

2. Beta Distribution

Simulation 1 for the Beta distribution is conducted in four steps. Step I consists of five cycles which are shown in Table 2. For each cycle, 5000 samples of size 5 are independently generated from each of the two Erlang distributions using fixed values of the parameters as shown in Table 2. These distributions are referred to as $\gamma_1(K_1,\lambda)$ and $\gamma_2(K_2,\lambda)$.

TABLE 2

Cycle Number	Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5
Values of Parameters $\gamma_1(K_1,\lambda)$	$K_1 = 4$ $\lambda = 3$	$K_1 = 4$ $\lambda = 3$	$K_1 = 4$ $\lambda = 3$	$K_1 = 4$ $\lambda = 3$	$K_1 = 4$ $\lambda = 3$
Values of Parameters $\gamma_2(K_2,\lambda)$	$K_2 = 4$ $\lambda = 3$	$K_2 = 5$ $\lambda = 3$	$K_2 = 6$ $\lambda = 3$	$K_2 = 7$ $\lambda = 3$	$K_2 = 8$ $\lambda = 3$
Mean (Beta)	.5	.44	.40	.36	.33
STD. DEV. (Beta)	.167	.1589	.150	.140	.1329

To illustrate the use of Table 2, 5000 samples of size 5 are generated for $\gamma_1(K_1,\lambda)$ distribution using the parameters $K_1 = 4$ and $\lambda = 3$. Another 5000 samples of size 5 are independently generated for $\gamma_2(K_2,\lambda)$ distribution using parameters $K_2 = 4$ and $\lambda = 3$ as shown in Table 2. During Cycle 2, 5000 samples of size 5 for each of the $\gamma_1(K_1,\lambda)$ and $\gamma_2(K_2,\lambda)$ distributions are independently generated using the parameters $K_1 = 4$, $\lambda = 3$ and $K_2 = 5$, $\lambda = 3$ respectively. These values are found in Table 2 under the heading Cycle 2. A similar process is repeated for the remaining Cycles 3, 4, and 5.

For each cycle, the $\gamma_1(K_1,\lambda)$ and $\gamma_2(K_2,\lambda)$ distributions are combined using the relationship $\gamma_1(K_1,\lambda)/\{\gamma_1(K_1,\lambda)+\gamma_2(K_2,\lambda)\}$ to form five 5000 samples of size 5 for the Beta distribution having mean $= \dfrac{K_1}{K_1+K_2}$ and variance $\dfrac{K_1\ K_2}{(K_1+K_2)^2(K_1+K_2+1)}$.

The actual values of these calculations for each cycle are shown in Table 2. It should be apparent that each cycle generates 5000 samples of size 5 from a different Beta distribution.

The remaining Steps II, III, and IV of this simulation are identical with Steps II, III, and IV of Simulation 1 under the Erlang distribution. Simulations 2, 3, and 4 for the Beta distribution use exactly the same techniques as outlined in Simulation 1 except that the sample sizes are 10, 15, and 20 respectively.

In addition, the effects of changes in the variance of the Beta distribution, when the mean is kept constant, are also investigated. This is accomplished by manipulating the values of the $\gamma_1(K_1,\lambda)$ and $\gamma_2(K_2,\lambda)$ parameters that result in Beta distributions having the same mean but different variances. The sets of values of these parameters used during this simulation are furnished in Table 3 below.

TABLE 3

Cycle #	Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5
Parameters of $\gamma_1(K_1,\lambda)$	$K_1 = 2$ $\lambda = 3$	$K_1 = 4$ $\lambda = 3$	$K_1 = 6$ $\lambda = 3$	$K_1 = 8$ $\lambda = 3$	$K_1 = 10$ $\lambda = 3$
Parameters of $\gamma_2(K_2,\lambda)$	$K_2 = 2$ $\lambda = 3$	$K_2 = 4$ $\lambda = 3$	$K_2 = 6$ $\lambda = 3$	$K_2 = 8$ $\lambda = 3$	$K_2 = 10$ $\lambda = 3$
Mean (Beta)	.5	.5	.5	.5	.5
STD. DEV. (Beta)	.036	.022	.0195	.013	.011

IV. RESULTS

The results of the various simulations for the Erlang and Beta distributions are furnished in tabular form in appendices C, D, E, F, G, and H. Appendix C contains Tables 4, 5, 6, and 7. These tables contain the frequency comparisons of the dis tributions obtained in these simulations and from Standard Normal designated N(0,1), when samples of size, 5, 10, 15, and 20 are used respectively. Tables 8, 9, 10, and 11 in Appendix D contain similar results obtained from the use of Beta distributions. Appendix E contains Tables 12 and 13. These tables have the comparison of the frequency distributions from N(0,1) and the transformed Beta(0,1) when the mean of the distribution is kept constant, and its variance is manipulated.

Appendices F, G, and H contain five tables each. These tables provide the necessary comparisons between the actual and observed significance levels for a fixed mean and 4, 9, 14, and 19 degrees of freedom. Tables 14, 15, 16, 17, and 18 in Appendix F pertain to Erlang distribution and contain the comparisons of the actual and the observed significance levels when samples of size 5, 10, 15, and 20 are used respectively. Tables 19, 20, 21, 22, and 23 in Appendix G contain similar results for the Beta distributions. Tables 24, 25, 26 27, and 28 in Appendix H have the comparisons of the significance levels when the variance of the Beta distribution is manipulated for a fixed mean.

Since there are but two different table formats used to illustrate the results, it is considered appropriate to discuss the use of only one table of each type before proceeding with

the discussion of the results. Tables of the first type, which give frequency comparisons, are used in Appendices C, D, and E. The second type of table is used to provide the comparisons of sigificance levels, and these tables are contained in Appendices F, G, and H.

As mentioned earlier, Table 4 pertains to the Erlang distribution. This table is divided into two parts. Part I furnishes the means and the standard deviations of the sampling distributions when the samples of size 5 are used. The means and standard deviations for each distribution are identified by means of index numbers. For example, Index (1) pertains to the Erlang distribution with mean 1.33 and standard deviation .6669. Part II of this table furnishes the frequency distributions of each Erlang given in Part I and the frequency distributions of N(0,1). In order to relate the mean and variance with the corresponding frequency distributions, the appropriate index should be consulted. For example, if one wishes to compare the frequency distribution of an Erlang with mean 2.0 and standard deviation .82, when using samples of size 5; one should consult Appendix C, Table 4. Corresponding to the above values, one would select Index 3 from Part I of Table 4 and obtain the frequency distribution from the row (3) in Part II of this table. In order to compare this frequency distribution with N(0,1), one would compare the corresponding values of row 3 with the row marked N(o,1). The asterisks on the distribution frequencies indicate that the distribution

was found to be significantly different than N(0,1) by chi-square goodness-of-fit test at 95% confidence level.

The second type of table is found in Appendices F, G, and H. Table 15 in Appendix F pertains to the Erlang distribution with a fixed mean of 1.667. This table provides the comparison of the actual significance levels and the observed significance levels when the sampling population is Erlang with mean 1.667. This table furnishes these comparisons for four values of degrees of freedom. These values are 4, 9, 14, and 19; they are found in the first column of the table. The format of this table is very similar to the standard t-table found in most text books. In order to make use of this table, one is required to know the sample size and the population mean. For example, if one is interested in determining the comparison of significance levels; when sampling from Erlang distribution with mean 2.0 using sample size 15, one would consult Appendix F, Table 16 and obtain the required information in the row marked 14. The asterisks on the observed significance values indicate that the results were found to be significantly different at 95% confidence level by the use of chi-square goodness-of-fit test.

RESULTS OF ERLANG

The results shown in Tables 4, 5, 6, and 7 indicate that the Erlang frequency distributions are significantly different than the N(0,1) at 5% level. A comparison of the observed significance levels with the actual significance levels in

Table 14, Appendix F reveals that when an Erlang distribution with mean 1.333 is used the observed alpha values are significantly different from the actual alpha values. For example, for Sample Size 5, corresponding to the actual values of α = .5, .2, and .05; the observed values are α = .034, .004, and .0004 respectively. It is further noted that as the sample size is increased to 10, 15, and 20, the corresponding observed values are zero in each case. This means that all the absolute values of t statistic are smaller than the critical values.

From the results in Table 15, one observes that for sample size 5 at α = .5, .2, .05, the observed values are α = .0292, .002, and .0002 respectively. Each of these values is smaller than the corresponding values in Table 14. The results for sample size 10, 15, and 20 are the same as found in Table 14. The results in Tables 16 and 17 indicate that for sample size 5, the corresponding observed values in each case are α = .0292, .0034, .0002 and .0268, .0028, .0002 respectively. The alpha values for sample sizes 10, 15, and 20 remain zero.

In examining the graphs of these distributions, one observes that as the mean of the distribution is increased; the distribution, though flatter than the normal distribution, becomes more nearly symmetric around the mean. However as the mean of the distribution is increased, the variance ($\frac{K}{\lambda^2}$) is also increased. Consequently, when sampling from this distribution, the absolute values of the t statistic, $\frac{\bar{x}-\mu}{s/\sqrt{N}}$, are

smaller than the absolute values of this statistic when sampling from the normal population. Furthermore, these values decrease as the sample size is increased.

From the foregoing results, it is concluded that "Student's" t-test is sensitive to the Erlang distribution. It is further observed that when sampling from an Erlang distribution, the resulting t statistic has significantly shorter tails than the "Student's" t distribution. It would seem, therefore, that the resulting t statistic is a form of leptokurtic curve.

RESULTS OF BETA DISTRIBUTION

Results in Tables 8, 9, 10, and 11 reveal that except when the sample size is 20; and the distribution is symmetric around the mean (.5), the Beta distribution frequencies are significantly different from the N(0,1). A comparison of the results in Table 19, Appendix G reveals that when sampling from the Beta distribution with mean .5, the observed alpha values in most cases are not significantly different from the actual alpha values. For example, for sample size 5, the observed alpha values are only significant at α = .8 and .5. However, when the sample size is increased to 10, 15, and 20; the results are only significant at α = .8. This suggests that when sampling from a symmetric Beta distribution, using samples of size 5 or greater, the "Student's" t-test yields results that are comparable with results obtained from normal sampling distributions.

The results in Table 20 reveal that when sampling from the Beta distribution with mean .44, the observed alpha values

for sample size 10 remain unchanged. The observe alpha values using sample size 5 are significantly different at $\alpha = .8$, .5, and .2. The observed alpha values for sample size 20 are significant at $\alpha = .8$, .5, .1 and .05. Comparing similar results in Tables 21 and 22, one observes that as the mean of the sampling distribution is varied to the left of the mean, the use of larger sample sizes (10, 15) yield observed alpha values that are comparable to the actual alpha values. In Table 23, one observes that when the sampling distribution has mean .33; i.e., it is significantly skewed to the right, the observed alpha values for sample size 5 are significantly different than the actual alpha values. The results for sample size 10 are reasonably comparable with the actual alpha values.

From the foregoing, it is concluded that the "Student's" t-test is sensitive to changes in the shape resulting from changes in the mean of the distribution. When the Beta distribution is symmetric around the mean, the use of "Student's" t-test yields results that are comparable with those obtained when sampling from the normal population. It is further noted that most of the observed alpha values are smaller than the actual alpha values. This suggests that the values of the t statistic are smaller than the corresponding values of "Student's" t statistic. It would seem that the tails of the new t statistic are shorter than the tails of "Student's" t.

The results of the Beta distributions with fixed mean and different variances are contained in Tables 24, 25, 26,

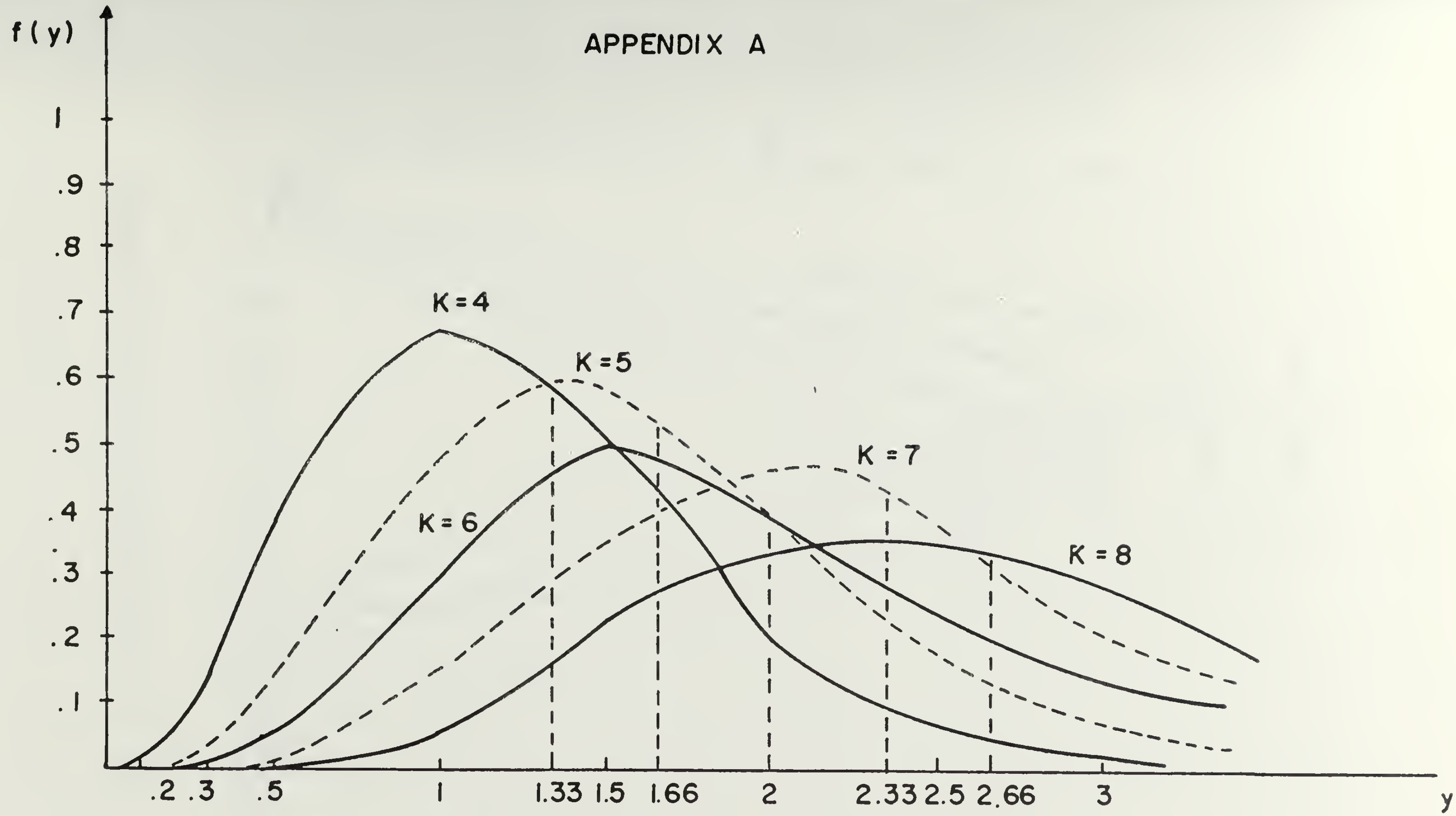
27, and 28. A similar investigation of these results indicates that for small sample size (5) the observed alpha values are significantly different than the true values. However, as the sample size is increased to 10, the difference between the observed and actual alpha values is significantly reduced. This suggests that "Student's" t-test is sensitive to changes in the variance when small samples are used.

V. CONCLUSIONS AND RECOMMENDATIONS

The simulation described on the foregoing pages which investigated the robustness of "Student's" t-test has proven to be extremely realistic. The results of violating the assumption of normality by the use of the Beta and the Gamma distributions indicate that when the sampling distribution is symmetric around the mean, the use of "Student's" t-test gives results that are closely comparable with the values obtained by the use of normal distributions. This implies that "Student's" t-test is reasonably robust when it is used with symmetric populations. The results also indicate that in the case of skewed distributions, the observed alpha values are significantly different from the actual significance levels. It is further noted that in the case of symmetric distributions, the values of observed significane levels approach the true values as the sample size is increased.

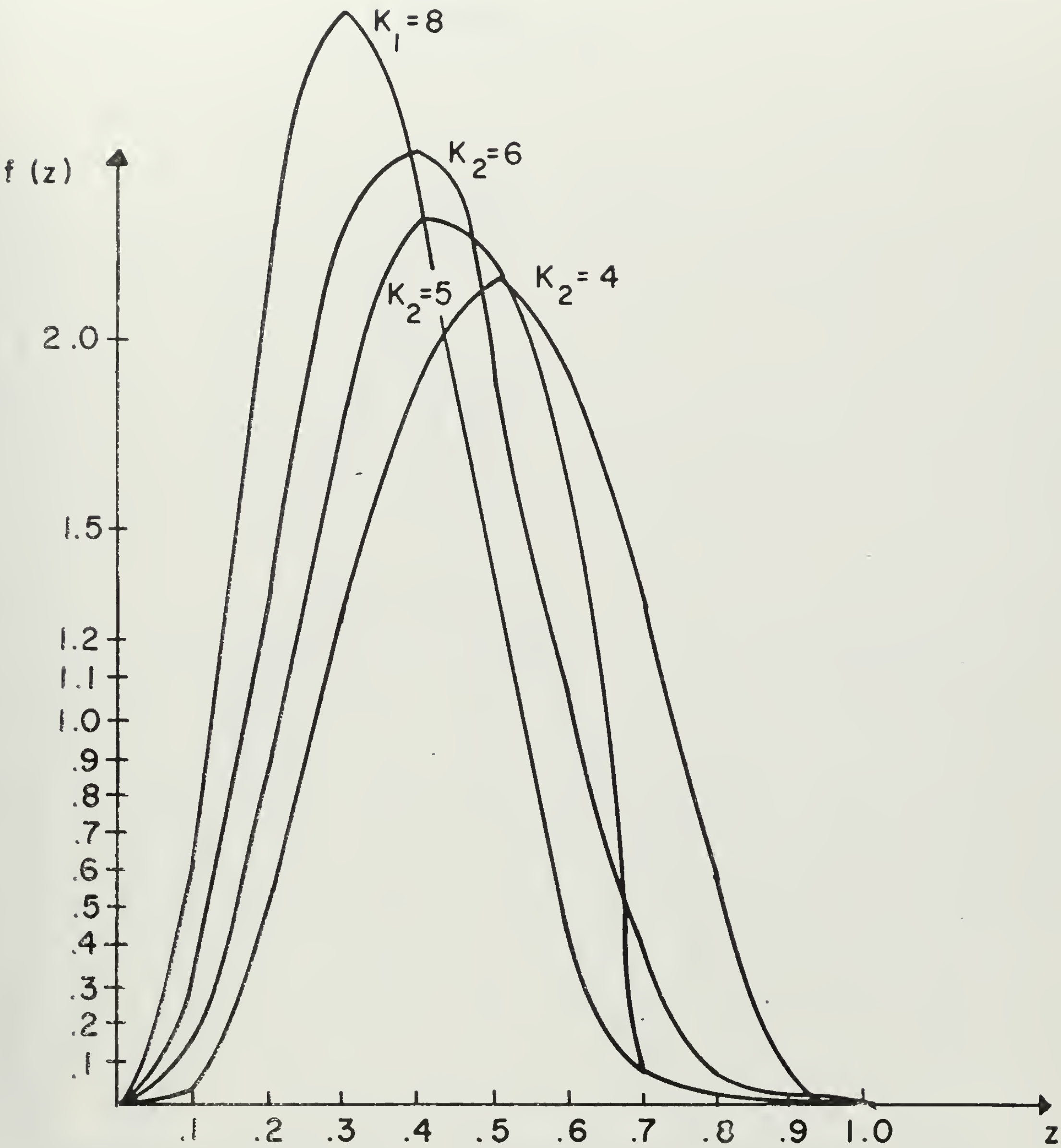
In the case of symmetrical distributions with a fixed mean and different variances, it is observed that the "Student's" t-test is sensitive to the change in the variance. The observed significance levels are considerably different from the true values as the variance of the sampling population is increased. However, as the sample size is increased, the effects of change in the variance are considerably reduced.

APPENDIX A



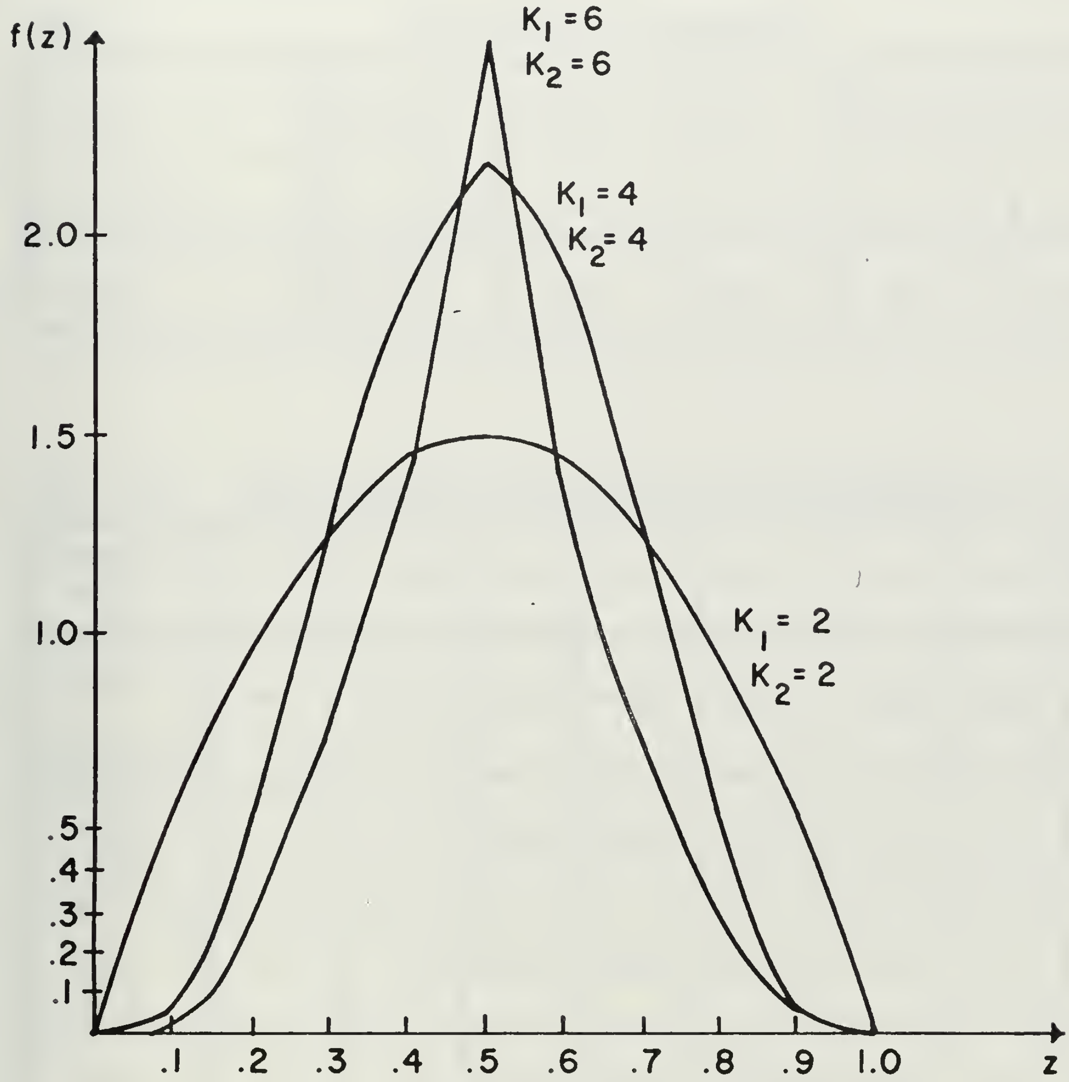
GRAPHS OF ERLANG DISTRIBUTIONS FOR $\lambda = 3$

APPENDIX B



GRAPH OF BETA DISTRIBUTIONS $\alpha = K_1 = 4$

APPENDIX B_1



GRAPH OF BETA DISTRIBUTIONS WITH MEAN = 0.5, DIFFERENT VARIANCE.

APPENDIX C

Table 4

POPULATION Erlang

SAMPLE SIZE 5

Part I Population Parameters

Mean	1.33	1.667	2.0	2.333	2.667
STD.DEV.	.6669	.74	.82	.86	.93
Index	(1)	(2)	(3)	(4)	(5)

Part II Comparisons of Distribution Frequencies

Inter-vals	< $\mu-3\sigma$	$\mu-\sigma$	$\mu-\sigma$	$\mu-\sigma$	$\mu+\sigma$	$\mu+\sigma$	$\mu+\sigma$	> $\mu+3\sigma$
N(0,1) γ(0,1)	74	524	3399	8499	8499	3399	524	74
Index (1)*	0	0	3581	10575	6991	2793	813	247
(2)*	0	10	3653	10250	7244	2820	990	233
(3)*	0	26	3699	10161	7262	2841	774	237
(4)*	0	62	3797	9743	7469	2948	802	179
(5)*	0	75	3824	9762	7390	2490	782	177

APPENDIX C

Table 5

POPULATION: Erlang

SAMPLE SIZE: 10

Part I Population Parameters

Mean	1.333	1.667	2.0	2.33	2.667
STD.DEV.	.6604	.749	.83	.905	.9755
Index	(1)	(2)	(3)	(4)	(5)

Part II Comparison of Distribution Frequencies

Intervals	< $\mu-3\sigma$	$\mu-3\sigma$	$\mu-2\sigma$	$\mu-\sigma$	$\mu+\sigma$	$\mu+2\sigma$	$\mu+3\sigma$	> $\mu+3\sigma$
N(0,1) γ(0,1)	149	1049	6799	16999	16999	6799	1049	149
Index (1)*	0	1	7294	20977	14043	5495	1670	520
(2)*	0	11	7029	20628	14691	5610	1572	459
(3)*	0	39	6854	20464	15090	5636	1500	417
(4)*	0	46	6702	20443	15197	5819	1447	346
(5)*	0	78	6670	20394	15226	5822	1490	320

APPENDIX C

Table 6

POPULATION: Erlang

SAMPLE SIZE: 15

Part I Population Parameters

Mean	1.33	1.667	2.0	2.333	2.667
STD.DEV.	.6725	.7677	.8466	.9173	.9983
Index	(1)	(2)	(3)	(4)	(5)

Part II Comparisons of Distribution Frequencies

Inter-vals	< $\mu-3\sigma$	$\mu-3\sigma$	$\mu-2\sigma$	$\mu-\sigma$	$\mu+\sigma$	$\mu+2\sigma$	$\mu+3\sigma$	> $\mu+3\sigma$
N(0,1) γ(0,1)	224	1574	10199	25499	25499	10199	1574	224
Index (1)*	0	0	10018	31845	21764	8223	2423	724
(2)*	0	4	9739	31562	22509	8283	2216	687
(3)*	0	25	9802	31101	22752	8495	2272	553
(4)*	0	54	9750	30661	23129	8783	2139	484
(5)*	0	75	9763	30455	23504	8666	2082	455

APPENDIX C

Table 7

POPULATION: Erlang

SAMPLE SIZE: 20

Part I Population Parameters

Mean	1.33	1.667	2.0	2.333	2.667
STD.DEV.	.6825	.7710	.8536	.9232	.9957
Index	(1)	(2)	(3)	(4)	(5)

Part II Comparisons of Distribution Frequencies

Inter-vals	< $\mu-3\sigma$	$\mu-3\sigma$	$\mu-2\sigma$	$\mu-\sigma$	$\mu+\sigma$	$\mu+2\sigma$	$\mu+3\sigma$	> $\mu+3\sigma$
N(0,1) γ(0,1)	299	2099	13599	33999	33999	13599	2099	299
Index (1)*	0	0	12775	43055	29172	11113	3142	943
(2)	0	5	12569	42147	30198	11222	3055	809
(3)*	0	30	12625	41768	30649	11277	2897	754
(4)*	0	69	12582	41276	31091	11553	2798	631
(5)*	0	109	12587	40938	31459	11636	2705	566

APPENDIX D

Table 8

POPULATION: Beta

SAMPLE SIZE: 5

Part I Population Parameters

Mean	.5	.44	.40	.36	.33
STD.DEV.	.167	.1589	.50	.140	.1329
Index	(1)	(2)	(3)	(4)	(5)

Part II Comparison of Distribution Frequencies

Interval	< $\mu-3\sigma$	$\mu-3\sigma$	$\mu-2\sigma$	$\mu-\sigma$	$\mu+\sigma$	$\mu+2\sigma$	$\mu+3\sigma$	> $\mu+3\sigma$
N(0,1)	74	524	3399	8499	8499	3399	524	74
Beta(0,1)								
(1)*	0	411	3879	8062	8333	3896	419	0
(2)*	0	298	3894	8414	8146	3670	575	3
(3)*	0	194	3890	8739	7987	3515	660	15
(4)*	0	151	3942	8919	7849	3445	656	38
(5)*	0	100	3919	9111	7736	3382	696	56

APPENDIX D

Table 9

POPULATION: Beta

SAMPLE SIZE: 10

Part I Population Parameters

Mean	.5	.44	.40	.36	.33
STD.DEV.	.167	.157	.148	.1397	.132
Index	(1)	(2)	(3)	(4)	(5)

Part II Comparison of Distribution Frequencies

Inter-val	< $\mu-3\sigma$	$\mu-3\sigma$	$\mu-2\sigma$	$\mu-\sigma$	$\mu+\sigma$	$\mu+2\sigma$	$\mu+3\sigma$	> $\mu+3\sigma$
N(o,1) Beta(0,1)	149	1049	6799	16999	16999	6799	1049	149
Index (1)*	0	802	7637	16467	16467	7726	935	0
(2)*	0	629	7841	16813	16123	7418	1164	10
(3)*	0	481	7796	14386	15766	7139	1387	45
(4)*	0	325	7879	17801	15558	6916	1457	64
(5)*	0	261	7789	18124	15378	6931	1396	121

APPENDIX D

Table 10

POPULATION: Beta

SAMPLE SIZE: 15

Part I Population Parameters

Mean	.5	.44	.4	.36	.33
STD.DEV.	.166	.1577	.1477	.1385	.1308
Index	(1)	(2)	(3)	(4)	(5)

Part II Comparisons of Distribution Frequencies

Inter-val	< $\mu-3\sigma$	$\mu-3\sigma$	$\mu-2\sigma$	$\mu-\sigma$	$\mu+\sigma$	$\mu+2\sigma$	$\mu+3\sigma$	> $\mu+3\sigma$
N(0,1) Beta(0,1)	224	1574	10199	25499	25499	10199	1574	224
Index (1)*	0	1225	11574	24441	24621	11812	1327	0
(2)*	0	931	11741	25404	23774	11407	1727	16
(3)*	0	683	11820	26111	23470	10762	2105	49
(4)*	0	530	11927	26398	23346	10494	2189	116
(5)*	0	429	11841	26916	23037	10283	2275	176

APPENDIX D

Table 11

POPULATION: Beta

SAMPLE SIZE: 20

Part I Population Parameters

Mean	.5	.44	.40	.36	.33
STD.DEV.	.167	.156	.1469	.138	.1297
Index	(1)	(2)	(3)	(4)	(5)

Part II Comparison of Distribution Frequencies

Inter-vals	< $\mu-3\sigma$	$\mu-3\sigma$	$\mu-2\sigma$	$\mu-\sigma$	$\mu+\sigma$	$\mu+2\sigma$	$\mu+3\sigma$	> $\mu+3\sigma$
N(0,1) Beta(0,1)	299	2099	13599	33999	33999	13599	2099	299
Index (1)*	0	1641	15458	32592	3271	15798	1780	0
(2)*	0	1252	15714	33539	32161	14858	2447	29
(3)*	0	970	15740	34596	31365	14505	2721	103
(4)*	0	766	15774	35284	31087	13905	3005	179
(5)*	0	586	15896	35825	30499	13853	3094	247

APPENDIX E

Table 12

POPULATION: Beta

SAMPLE SIZE: 5 (Fixed Mean, Difference Variance)

Part I Population Parameters

Mean	.5	.5	.5	.5	.5
Variance	.1036	.022	.0195	.013	.0108
Index	(1)	(2)	(3)	(4)	(5)

Part II Comparison of Distribution Frequencies

Interval	< $\mu-3\sigma$	$\mu-3\sigma$	$\mu-2\sigma$	$\mu-\sigma$	$\mu+\sigma$	$\mu+2\sigma$	$\mu+3\sigma$	> $\mu+3\sigma$
N(0,1) Beta(0,1)	74	524	3399	8499	8499	3399	524	74
Index (1)*	0	166	4394	7941	7764	4572	165	0
(2)*	0	397	3908	8075	8336	3846	436	0
(3)*	3	430	3584	8573	8456	3659	493	2
(4)*	3	439	3573	8399	8492	3597	486	11
(5)*	9	431	3456	8496	8575	3581	439	11

APPENDIX E

Table 13

POPULATION: Beta

SAMPLE SIZE: 10 (Fixed Mean, Difference Variance)

Part I Population Parameters

Mean	.5	.5	.5	.5	.5
Variance	.036	.022	.0195	.013	.0108
Index	(1)	(2)	(3)	(4)	(5)

Part II Comparison of Distribution Frequencies

Inter-vals	< $\mu-3\sigma$	$\mu-3\sigma$	$\mu-2\sigma$	$\mu-\sigma$	$\mu+\sigma$	$\mu+2\sigma$	$\mu+3\sigma$	> $\mu+3\sigma$
N(0,1)	149	1049	6799	16999	16999	6799	1049	149
Beta(0,1) Index (1)*	0	399	8788	15650	15758	8963	442	0
(2)*	0	838	7653	16251	16653	7811	894	0
(3)*	8	861	7361	16598	16669	7534	959	11
(4)*	15	946	7026	16749	16940	7318	985	21
(5)	12	919	6985	16834	16964	7226	1033	27

APPENDIX F

Table 14

POPULATION: Erlang

POPULATION MEAN: 1.33

Degrees of Freedom	VALUES OF 1-α							
	.2	.5	.8	.9	.95	.98	.99	.999
4	.7416*	.9660*	.9960*	.9982*	.9996*	1.0*	1.0*	1.0
9	.9586*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0
14	.9946*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0
19	.9996*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0

APPENDIX F

Table 15

POPULATION: Erlang

POPULATION MEAN: 1.667

Degrees of Freedom	VALUES OF 1-α							
	.2	.5	.8	.9	.95	.98	.99	.999
4	.7480*	.9708*	.9980*	.9998*	.9998*	1.0*	1.0*	1.0
9	.9650*	.998*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0
14	.9952*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0
19	.9998*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0

APPENDIX F

Table 16

POPULATION: Erlang

POPULATION MEAN: 2.0

Degrees of Freedom	VALUES OF 1-α							
	.2	.5	.8	.9	.95	.98	.99	.999
4	.7466*	.9708*	.9966*	.9994*	.9998*	1.0*	1.0*	1.0
9	.9608	.999*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0
14	.9958*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0*
19	1.0*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0

APPENDIX F

Table 17

POPULATION: Erlang

POPULATION MEAN: 2.33

Degrees of Freedom	VALUES OF 1-α							
	.2	.5	.8	.9	.95	.98	.99	.999
4	.7394*	.9732*	.9972*	.9994*	.9998*	1.0*	1.0*	1.0
9	.9640*	.999*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0
14	.9970*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0
19	1.0*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0

APPENDIX F

Table 18

POPULATION: Erlang

POPULATION MEAN: 2.667

Degrees of Freedom	VALUES OF 1-α							
	.2	.5	.8	.9	.95	.98	.99	.999
4	.754*	.9750*	.9982*	.998*	1.0*	1.0*	1.0*	1.0
9	.9656*	.999*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0
14	.9962*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0
19	1.0*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0*	1.0

APPENDIX G

Table 19

POPULATION: BETA

POPULATION MEAN: .5

Degrees of Freedom	ACTUAL SIGNIFICANCE LEVELS							
	.8	.5	.2	.1	.05	.02	.01	.001
4	.793*	.485*	.196	.099	.051	.020	.011	.002
9	.795*	.497	.199	.099	.051	.023	.014	.002
14	.807*	.498	.196	.095	.050	.040	.010	.001
19	.809*	.501	.199	.102	.051	.021	.01	.001

APPENDIX G

Table 20

POPULATION: Beta

POPULATION MEAN: .44

Degrees of Freedom	ACTUAL SIGNIFICANCE LEVELS							
	.8	.5	.2	.1	.05	.02	.01	.001
4	.793*	.480*	.190*	.099	.050	.021	.001	.001
9	.805*	.053	.200	.102	.050	.021	.011	.001
14	.799	.489*	.193*	.098	.050	.041	.012	.001
19	.805*	.509*	.200	.109*	.058*	.023	.02	.001

APPENDIX G

Table 21

POPULATION: Beta

POPULATION MEAN: .40

Degrees of Freedom	ACTUAL SIGNIFICANCE LEVELS							
	.8	.5	.2	.1	.05	.02	.01	.001
4	.793*	.488*	.192*	.096	.050	.020	.010	.001
9	.802	.491*	.192*	.096	.048	.021	.010	.001
14	.803*	.503	.196	.101	.053	.043	.011	.002
19	.791*	.494*	.189	.094	.046	.018	.01	.002

APPENDIX G

Table 22

POPULATION: Beta

POPULATION MEAN: .36

Degrees of Freedom	ACTUAL SIGNIFICANCE LEVELS							
	.8	.5	.2	.1	.05	.02	.01	.001
4	.798	.492*	.204	.098	.053	.020	.010	.001
9	.793*	.494*	.194	.102	.054	.021	.012	.001
14	.795*	.507*	.204	.100	.051	.043	.011	.002
19	.803*	.501	.196	.094	.050	.022	.01	.001

APPENDIX G

Table 23

POPULATION: Beta

POPULATION MEAN: .33

Degrees of Freedom	ACTUAL SIGNIFICANCE LEVELS							
	.8	.5	.2	.1	.05	.02	.01	.001
4	.801	.504	.205	.107*	.061*	.023*	.012*	.003*
9	.793*	.504	.200	.101	.052	.023	.011	.001
14	.796*	.507*	.209*	.106	.052	.046	.013	.002
19	.794*	.493*	.193*	.103	.0[illegible]4	.023	.02	.001

APPENDIX H

Table 24

POPULATION: Beta (Fixed Mean- Different Variance)

POPULATION MEAN: .5

POPULATION VARIANCE: .036

Degrees of Freedom	ACTUAL SIGNIFICANCE LEVELS							
	.8	.5	.2	.1	.05	.02	.01	.001
4	.7860*	.4796*	.2084*	.1160*	.0682*	.0314*	.0176*	.0012
9	.7952*	.4870*	.1948	.1054	.0562*	.0244	.0132	.0018

APPENDIX H

Table 25

POPULATION: Beta (Fixed Mean-Different Variance)

POPULATION MEAN: .5

POPULATION VARIANCE: .022

Degrees of Freedom	ACTUAL SIGNIFICANCE LEVELS							
	.8	.5	.2	.1	.05	.02	.01	.001
4	.7874*	.4842*	.2002	.1052	.0538	.0238	.0118	.0012
9	.8008	.4934	.1992	.0998	.0538	.0220	.0130	.0012

APPENDIX H

Table 26

POPULATION: Beta (Fixed Mean-Different Variance)

POPULATION MEAN: .5

POPULATION VARIANCE: .0195

Degrees of Freedom	ACTUAL SIGNIFICANCE LEVELS							
	.8	.5	.2	.1	.05	.02	.01	.001
4	.7908*	.4914*	.1948	.1068	.0556	.0224	.0102	.0008
9	.8012	.5042	.1956	.0976	.0486	.0220	.0094	.0008

APPENDIX H

Table 27

POPULATION: Beta

POPULATION MEAN: .5

POPULATION VARIANCE: .013

Degrees of Freedom	ACTUAL SIGNIFICANCE LEVELS							
	.8	.5	.2	.1	.05	.02	.01	.001
4	.7864*	.4824*	.1868*	.0924*	.0484	.0188	.0084	.0002
9	.7910*	.4908*	.1940	.0970	.0492	.0196	.0112	.0018

APPENDIX H

Table 28

POPULATION: Beta

POPULATION MEAN: .5

POPULATION VARIANCE: .0108

Degrees of Freedom	ACTUAL SIGNIFICANCE LEVELS							
	.8	.5	.2	.1	.05	.02	.01	.001
4	.8026*	.4906*	.2004	.1	.0498	.0204	.0104	.0014
9	.7910	.5030	.2010	.0974	.0464	.0172	.0090	.0008

APPENDIX J

Instructions for the Use of Computer Program

The attached computer program was used to obtain the simulation results for the Beta distribution. The program, written in Fortran language, makes use of an IBM 360 computer. This program consists of four simulations, requires 47 minutes of execution time, and makes use of 505 K storage capacity. The program as presented in Appendix J, is self sufficient in that it does not require any external sub-routines or programs for its execution. The program contains its own random number generator with the variant seed characteristics. It enables the user to manipulate the various parameters without any reorganization of the program.

To assist users in understanding and utilizing various facets of the program, a number of comment cards are used through-out the program. In order to illustrate the techniques of varying the parameters of the Erlang distribution or the number of cycles in the simulation, or the number of samples for the distributions, the following instructions could be useful. If one consults Appendix J, page 62, he should find that parameters $B_1 = 3$ corresponds to $\lambda = 3$, and the values of K_1 and K_2 represent the values of parameter K in the Erlang distribution. In order to obtain five different Erlang distributions in each cycle of each simulation (as illustrated in Table 2), one would set $K_1 = 4$, $K_2 = 3$ and make use of the Do loop 777. Thus in order to change the number of cycles to 3, one would simply replace Do 777 IB=1,5 with Do 777 IB=1,3.

This program is designed to furnish 4 different sample sizes. This is accomplished by the first Do loop on page 62; i.e., Do 778 IN=1,4. The sample size is determined by the value of the increment in statement, MR = MR + 5. If four samples of sample size 2, 4, 6, and 8 are required, one would simply change the statement MR = MR + 2. If only two samples of 5000 each are desired, then one would replace Do 778 IN = 1,4 by Do 778 IN = 1,2. The number of samples in each simulation are determined by the Do loop "Do 1000 II = 1,50. As is apparent from the dimensions of GAMMA1 and GAMMA2, the program for given values of K_1 and K_2 generates 100 samples of size MR which are Beta distributed. It calculates the new t statistic, compares the absolute values of this statistic with the critical values, and stores the beta variate in location called "STOR". Utilizing a different seed, the entire process is repeated 50 times, thus yielding a total of 5000 samples of size MR. After the generation of 5000 samples which are located in "STOR", the distribution frequency comparisons are performed; and the desired results are provided as indicated in Appendix J. The values of parameters are then incremented by the Do loop 778, and the entire process is repeated until 5 cycles are completed for the value of sample size. At the end of Simulation I, the sample size is incremented, and the entire process is repeated.

A sample of the results of one simulation are furnished in Appendix J. It may be recalled that each simulation

consists of five cycles; for each cycle, the program furnishes the following information as shown in Appendix J:

1. Transformed Beta frequency distribution denoted by (1) .
2. N(0,1) frequency distribution denoted by (2) .
3. Sample size denoted by (3) .
4. Distribution mean denoted by (4) .
5. The actual levels of significance denoted by (5) .
6. The observed levels of significance denoted by (6) .
7. The $X^2_{.95}(1)$ statistic for each level of significance denoted by (7) .

At the end of the simulation, a summary of the results is furnished which contains the actual alpha values, the observed alpha values for each of the five cycles, $X^2_{.95}$ values for one degree of freedom, and the observed $X^2(1)$ statistics for each significance level. The computer program for the Gamma distribution also utilizes the same techniques. This program is furnished in Appendix J.

COMPUTER PROGRAM FOR ERLANG DISTRIBUTION

```
      REAL*8 GAMMA1(100,20),GAMMA2(100,20),GAMMA3(100,20),RIZWAN(100,2)/
     *200*0.0/,TOBS2(100),A,X
      COMMON (SMIN,SMAX,XM,XL,ILA1,ILA1A,ILA2,ILA2A,ILA3,ILA3A,ILA4,ILA4
     *A,ILA5,ILA5A,ILA6,ILA6A,ILA7,ILA7A,ILA8,ILA8A,S)
      INTEGER ACR(50),IPT(50,2)/100*0/,IM (50)/50*0/,IBLK/2H  /,SYM/2H *
     */ ,ACRA(100),SYMA/1H*/,INTVAL(8)/8*0/,GROUP(8)/8*0/,S
      INTEGER GR1,GR2,GR3,GTGR3,GL1,GL2,GL3,GTGL3,MR
      DIMENSION RIZ(100000),D(40),DA(40)
      B1=3.
      K1=0
      K2=0
      IX=53693
      K1=4
      MR=0
      DO 778 IN=1,4
      K2=3
      MR=MR+5
      NN=0
      DO 777 IB=1,5
      MMR=MR*5000
      CALL CLEAN
      K2=K2+1
      CD=(1.0*K1)/(K1+K2)
      DO 1000 II=1,50
      DO 907 IIA=1,50
  907 IM(IIA)=0
C    THE GENERATION OF GAMMA RANDOM VARIABLES IS AS FOLLOWS
      DO 10 I=1,100
      DO 10 J=1,MR
      YFL = 0.0
      DO 13 I=1,100
      DO 13 J=1,MR
      TR=0.0
      DO 12 JJ = 1,K2
      IY =  IX*65539
      IF(IY) 7,8,8
    7 IY=IY+2147483647+1
    8 YFL=IY
      YFL = YFL*.4656613E-9
      IX = IY
   12 TR=TR+ALOG(YFL)
      X=(-1.0/B1)*(TR)
      IF(X.LT.XL) XL=X
      IF(X.GT.XM) XM=X
   13 GAMMA2(I,J) =X
 9050 DO 905 I=1,100
      DO 905 J=1,MR
      S=S+1
```

```
  905 RIZ(S)=GAMMA2(I,J)                                                RIZ00490
      DO 250 I=1,100                                                    RIZ00500
      DO 250 J=1,MR                                                     RIZ00510
      RIZWAN(I,1) =RIZWAN(I,1)+ GAMMA2(I,J)                             RIZ00520
  250 RIZWAN(I,2) = RIZWAN(I,2)+ GAMMA2(I,J)**2                         RIZ00530
      DO 87 I=1,100                                                     RIZ00540
      RIZWAN(I,2) =DSQRT((RIZWAN(I,2)-RIZWAN(I,1)*RIZWAN(I,1)/MR)/(MR-1)RIZ00550
     *)                                                                 RIZ00560
   87 RIZWAN(I,1) =RIZWAN(I,1)/MR                                       RIZ00570
C     THE COMPARISON OF SIGNIFICANCE LEVELS IS AS FOLLOWS               RIZ00580
      IF(MR.EQ.5) GO TO 62                                              RIZ00590
      IF(MR.EQ.10) GO TO 63                                             RIZ00600
      IF(MR.EQ.15) GO TO 64                                             RIZ00610
      IF(MR.EQ.20) GO TO 65                                             RIZ00620
 62   RMN=K2/3.0                                                        RIZ00630
      DO 132 I=1,100                                                    RIZ00640
      TOBS2(I)=(RIZWAN(I,1)-RMN)/(RIZWAN(I,2))/MR**.5                   RIZ00650
      IF(DABS(TOBS2(I)).LE.0.271) ILA1=ILA1+1                           RIZ00660
      IF(DABS(TOBS2(I)).GT. 0.271)ILA1A=ILA1A+1                         RIZ00670
      IF(DABS(TOBS2(I)).LE.0.741) ILA2=ILA2+1                           RIZ00680
      IF(DABS(TOBS2(I)).GT.0.741) ILA2A=ILA2A+1                         RIZ00690
      IF(DABS (TOBS2(I)).LE.1.533) ILA3=ILA3+1                          RIZ00700
      IF(DABS(TOBS2(I)).GT.1.533) ILA3A=ILA3A+1                         RIZ00710
      IF(DABS (TOBS2(I)).LE.2.132) ILA4=ILA4+1                          RIZ00720
      IF(DABS (TOBS2(I)).GT.2.132) ILA4A=ILA4A+1                        RIZ00730
      IF(DABS (TOBS2(I)).LE.2.776) ILA5=ILA5+1                          RIZ00740
      IF(DABS (TOBS2(I)).GT. 2.776)ILA5A=ILA5A+1                        RIZ00750
      IF(DABS (TOBS2(I)).LE.3.747) ILA6=ILA6 +1                         RIZ00760
      IF(DABS(TOBS2(I)).GT.3.747)  ILA6=ILA6+1                          RIZ00770
      IF(DABS(TOBS2(I)).LE.4.604) ILA7=ILA7+1                           RIZ00780
      IF(DABS(TOBS2(I)).GT.4.604) ILA7A=ILA7A+1                         RIZ00790
      IF(DABS(TOBS2(I)).LE.8.610) ILA8=ILA8+1                           RIZ00800
      IF(DABS(TOBS2(I)).GT.8.610) ILA8A=ILA8A+1                         RIZ00810
  132 CONTINUE                                                          RIZ00820
      GO TO 888                                                         RIZ00830
 63   RMN=K2/3.0                                                        RIZ00840
      DO 133 I=1,100                                                    RIZ00850
      TOBS2(I)=(RIZWAN(I,1)-RMN)/(RIZWAN(I,2))/MR**.5                   RIZ00860
      IF(DABS(TOBS2(I)).LE.0.261) ILA1=ILA1+1                           RIZ00870
      IF(DABS(TOBS2(I)).GT. 0.261)ILA1A=ILA1A+1                         RIZ00880
      IF(DABS(TOBS2(I)).LE.0.703) ILA2=ILA2+1                           RIZ00890
      IF(DABS(TOBS2(I)).GT.0.703) ILA2A=ILA2A+1                         RIZ00900
      IF(DABS (TOBS2(I)).LE.1.383) ILA3=ILA3+1                          RIZ00910
      IF(DABS(TOBS2(I)).GT.1.383) ILA3A=ILA3A+1                         RIZ00920
      IF(DABS (TOBS2(I)).LE.1.833) ILA4=ILA4+1                          RIZ00930
      IF(DABS (TOBS2(I)).GT.1.833) ILA4A=ILA4A+1                        RIZ00940
      IF(DABS (TOBS2(I)).LE.2.262) ILA5=ILA5+1                          RIZ00950
      IF(DABS (TOBS2(I)).GT. 2.262)ILA5A=ILA5A+1                        RIZ00960
```

```
      IF(DABS (TOBS2(I)).LE.2.821) ILA6=ILA6 +1
      IF(DABS(TOBS2(I)).GT.2.821)  ILA6=ILA6+1
      IF(DABS(TOBS2(I)).LE.3.250) ILA7=ILA7+1
      IF(DABS(TOBS2(I)).GT.3.250) ILA7A=ILA7A+1
      IF(DABS(TOBS2(I)).LE.4.781) ILA8=ILA8+1
      IF(DABS(TOBS2(I)).GT.4.781) ILA8A=ILA8A+1
  133 CONTINUE
      GO TO 888
   64 RMN=K2/3.0
      DO 134 I=1,100
      TOBS2(I)=(RIZWAN(I,1)-RMN)/(RIZWAN(I,2))/MR**.5
      IF(DABS(TOBS2(I)).LE.0.258) ILA1=ILA1+1
      IF(DABS(TOBS2(I)).GT. 0.258)ILA1A=ILA1A+1
      IF(DABS(TOBS2(I)).LE.0.692) ILA2=ILA2+1
      IF(DABS(TOBS2(I)).GT.0.692) ILA2A=ILA2A+1
      IF(DABS (TOBS2(I)).LE.1.345) ILA3=ILA3+1
      IF(DABS(TOBS2(I)).GT.1.345) ILA3A=ILA3A+1
      IF(DABS (TOBS2(I)).LE.1.761) ILA4=ILA4+1
      IF(DABS (TOBS2(I)).GT.1.761) ILA4A=ILA4A+1
      IF(DABS (TOBS2(I)).LE.2.145) ILA5=ILA5+1
      IF(DABS (TOBS2(I)).GT. 2.145)ILA5A=ILA5A+1
      IF(DABS (TOBS2(I)).LE.2.264) ILA6=ILA6 +1
      IF(DABS(TOBS2(I)).GT.2.264) ILA6A=ILA6A+1
      IF(DABS(TOBS2(I)).LE.2.977) ILA7=ILA7+1
      IF(DABS(TOBS2(I)).GT.2.977) ILA7A=ILA7A+1
      IF(DABS(TOBS2(I)).LE.4.140) ILA8=ILA8+1
      IF(DABS(TOBS2(I)).GT.4.140) ILA8A=ILA8A+1
  134 CONTINUE
      GO TO 888
   65 RMN=K2/3.0
      DO 135 I=1,100
      TOBS2(I)=(RIZWAN(I,1)-RMN)/(RIZWAN(I,2))/MR**.5
      IF(DABS(TOBS2(I)).LE.0.257) ILA1=ILA1+1
      IF(DABS(TOBS2(I)).GT. 0.257)ILA1A=ILA1A+1
      IF(DABS(TOBS2(I)).LE.0.688) ILA2=ILA2+1
      IF(DABS(TOBS2(I)).GT.0.688) ILA2A=ILA2A+1
      IF(DABS (TOBS2(I)).LE.1.382) ILA3=ILA3+1
      IF(DABS(TOBS2(I)).GT.1.328) ILA3A=ILA3A+1
      IF(DABS (TOBS2(I)).LE.1.729) ILA4=ILA4+1
      IF(DABS (TOBS2(I)).GT.1.729) ILA4A=ILA4A+1
      IF(DABS (TOBS2(I)).LE.2.093) ILA5=ILA5+1
      IF(DABS (TOBS2(I)).GT. 2.093)ILA5A=ILA5A+1
      IF(DABS (TOBS2(I)).LE.2.539) ILA6=ILA6 +1
      IF(DABS(TOBS2(I)).GT.2.539)  ILA6=ILA6+1
      IF(DABS(TOBS2(I)).LE.2.861) ILA7=ILA7+1
      IF(DABS(TOBS2(I)).GT.2.861) ILA7A=ILA7A+1
      IF(DABS(TOBS2(I)).LE.3.883) ILA8=ILA8+1
      IF(DABS(TOBS2(I)).GT.3.883) ILA8A=ILA8A+1
```

```
135   CONTINUE                                                    RIZ01450
888   LOUIS=0.0                                                   RIZ01460
      DO 1000 I=1,100                                             RIZ01470
      RIZWAN(I,1)=0.0                                             RIZ01480
1000  RIZWAN(I,2)=0.0                                             RIZ01490
      Y1=ILA1/5000.                                               RIZ01500
      Y2=ILA2/5000.                                               RIZ01510
      Y3=ILA3/5000.                                               RIZ01520
      Y4=ILA4/5000.                                               RIZ01530
      Y5=ILA5/5000.                                               RIZ01540
      Y6=ILA6/5000.                                               RIZ01550
      Y7=ILA7/5000.                                               RIZ01560
      Y8=ILA8/5000.                                               RIZ01570
      WRITE(6,889) Y1,Y2,Y3,Y4,Y5,Y6,Y7,Y8                        RIZ01580
889   FORMAT(1H0,8F10.5)                                          RIZ01590
      A1=0.2                                                      RIZ01600
      A2=0.5                                                      RIZ01610
      A3=0.8                                                      RIZ01620
      A4=0.9                                                      RIZ01630
      A5=0.95                                                     RIZ01640
      A6=0.98                                                     RIZ01650
      A7=0.99                                                     RIZ01660
      A8=0.999                                                    RIZ01670
      CQ1=.098                                                    RIZ01680
      CQ2=.455                                                    RIZ01690
      CQ3=1.324                                                   RIZ01700
      CQ4=2.71                                                    RIZ01710
      CQ5=3.84                                                    RIZ01720
      CQ6=5.02                                                    RIZ01730
      CQ7=6.63                                                    RIZ01740
      CQ8=7.88                                                    RIZ01750
      F1=(ILA1-.2*5000)**2/(.2*5000)                              RIZ01760
      F2=(ILA1A-.8*5000)**2/(.8*5000)                             RIZ01770
      C1=F1+F2                                                    RIZ01780
      F3=(ILA2-.5*5000)**2/(.5*5000)                              RIZ01790
      F4=(ILA2A-.5*5000)**2/(.5*5000)                             RIZ01800
      C2=F3+F4                                                    RIZ01810
      F5=(ILA3-.8*5000)**2/(.8*5000)                              RIZ01820
      F6=(ILA3A-.2*5000)**2/(.2*5000)                             RIZ01830
      C3=F5+F6                                                    RIZ01840
      F7=(ILA4-.9*5000)**2/(.9*5000)                              RIZ01850
      F8=(ILA4A-.1*5000)**2/(.1*5000)                             RIZ01860
      C4=F7+F8                                                    RIZ01870
      F9=(ILA5-.95*5000)**2/(.95*5000)                            RIZ01880
      F10=(ILA5A-.05*5000)**2/(.05*5000)                          RIZ01890
      C5=F9+F10                                                   RIZ01900
      F11=(ILA6 -.98*5000)**2/(.98*5000)                          RIZ01910
      F12=(ILA6A-.02*5000)**2/(.02*5000)                          RIZ01920
```

```
      C6=F11+F12
      F13=(ILA7-.99*5000)**2/(.99*5000)
      F14=(ILA7A-.01*5000)**2/(.01*5000)
      C7=F13+F14
      F15=(ILA8 -.999*5000)**2/(.999*5000)
      F16=(ILA8A-.001*5000)**2/(.001*5000)
      C8=F15+F16
      D(NN+IB)=Y1
      D(NN+IB+1)=Y2
      D(NN+IB+2)=Y3
      D(NN+IB+3)=Y4
      D(NN+IB+4)=Y5
      D(NN+IB+5)=Y6
      D(NN+IB+6)=Y7
      D(NN+IB+7)=Y8
      DA(NN+IB)=C1
      DA(NN+IB+1)=C2
      DA(NN+IB+2)=C3
      DA(NN+IB+3)=C4
      DA(NN+IB+4)=C5
      DA(NN+IB+5)=C6
      DA(NN+IB+6)=C7
      DA(NN+IB+7)=C8
      NN=NN+7
C
C*****************************************************************************
C     THE COMPARISON OF FREQUENCY  INTERVALS IS AS FOLLOWS
      DO 111 I=1,8
 111  INTVAL(I)=0
      XBAR2= XBAR(RIZ,MMR)
      SIGMA2 = STDDVN(RIZ,MMR)
      BASE2=-3.0
      DO 72 I = 1,MMR
      IF((RIZ(I)-XBAR2)/SIGMA2 .LE.BASE2) GROUP(1)=GROUP(1)+1
      DO 73 J=2,7
      K=J-2
      IF((RIZ(I)-XBAR2)/SIGMA2 .GT.BASE2+K.AND.( RIZ(I)-XBAR2)/SIGMA2 .L
     *E.BASE2+(K+1)) GROUP(J)=GROUP(J)+1
  73  CONTINUE
      IF((RIZ(I) -XBAR2)/SIGMA2.GT.BASE2+6) GROUP(8)=GROUP(8)+1
  72  CONTINUE
C*****************************************************************************
C********************************************************************
      N=MMR
      GR1=(0.34)*N
      GR2=(0.136)*N
      GR3=(0.021)*N
      GTGR3=(0.003)*N
```

```
      GL1=(0.34)*N
      GL2=(0.136)*N
      GL3=(0.021)*N
      GTGL3=(0.003)*N
C*****************************************
C**********************************************************************
C   VARIOUS WRITE STATEMENTS ARE AS FOLLOWS
      WRITE(6,74)
      WRITE(6,75) GTGL3,GL3,GL2,GL1,GR1,GR2,GR3,GTGR3
      WRITE(6,75)GROUP
      WRITE(6,151) MR
      WRITE(6,553)
      WRITE(6,603) RMN
      WRITE(6,672)
      WRITE(6,671) SIGMA2
      WRITE(6,673)
 71   FORMAT('0','FOR GAMMA NUMBR TOBS LETCRIT AT ALPHA=    ',I5,' & NMBR
     *GT TCRIT=',I5,//)
 74   FORMAT(1H0,' CHARACTERISTICS OF GAMMA DIST.  USED ARE :',//)
  75  FORMAT(8I10)
 99   FORMAT('0',' GAMM1 RANDOM NUMBERS ARE AS FOLLOWS ')
  100 FORMAT('0',' GAMM2 RANDOM NUMBERS ARE AS FOLLOWS ')
 101  FORMAT('0', ' IX VALUE = ',I15,' YFL VALUE = ',F12.8)
 151  FORMAT(1H0,' SAMPLE SIZE = ',I5,//)
 553  FORMAT(1H1,'PAGE NUMBER =  SAMPLE MEAN = ',//)
 603  FORMAT(1H0,' DISTRIBUTION MEAN = ',F10.3,//)
 671  FORMAT(1H0,' STANDARD DEVIATION = ',F10.5,//)
 672  FORMAT(1H0,' FREQUENCY COUNTS FOR GAMMA DISTRIBUTION ARE ',//)
 673  FORMAT(1H0,' FREQUENCY COUNTS FOR  NORMAL DISTRIBUTION ARE ',//)
C**********************************************************************
C**********************************************************************
      DO 7009 I=1,100
 7009 TOBS2(I)=0
 777  CONTINUE
      WRITE(6,609)
 609  FORMAT(1H1,' COMPARISON OF   1 - ALPHA  VALUES IS AS FOLLOWS ',//)
      WRITE(6,607)
 607  FORMAT(1H0,' CRITICAL VALUES ',//)
      WRITE(6,604) A1,A2,A3,A4,A5,A6,A7,A8
 604  FORMAT(20X,8F10.5,//)
      WRITE(6,608)
 608  FORMAT(1H0,' EMPIRICAL VALUES ',//)
      WRITE(6,604)D
      WRITE(6,155)
  155 FORMAT(1H0,'CHISQUARE CRITICAL VALUES ARE',//)
      WRITE(6,604)CQ1,CQ2,CQ3,CQ4,CQ5,CQ6,CQ7,CQ8
      WRITE(6,156)
  156 FORMAT(1H0,'CHI SQ STATISTICS IS AS FOLLOWS',//)
```

```
      WRITE(6,604)DA
 778  CONTINUE
      WRITE(6,7001)
7001  FORMAT(1H1,20X,'END OF JOB')
      STOP
      END
C*********************************************************************

      SUBROUTINE CLEAN
      COMMON (SMIN,SMAX,XM,XL,ILA1,ILA1A,ILA2,ILA2A,ILA3,ILA3A,ILA4,IL A4
     *A,ILA5,ILA5A,ILA6,ILA6A,ILA7,ILA7A,ILA8,ILA8A,S)
      SMIN=0
      SMAX=0
      XM=0
      XL=0
      ILA1=0
      ILA2=0
      ILA3=0
      ILA4=0
      ILA5=0
      ILA6=0
      ILA7=0
      ILA8=0
      S=0
      ILA1A=0
      ILA2A=0
      ILA3A=0
      ILA4A=0
      ILA5A=0
      ILA6A=0
      ILA7A=0
      ILA8A=0
      RETURN
      END
C*********************************************************************
C
C*******************************************

      FUNCTION XBAR(X,N)
      DIMENSION X(1)
      SUMX=0.
      DO 100 I=1,N
 100  SUMX=SUMX+X(I)
      XBAR = SUMX/N
      RETURN
      END
```

```
C***************************************
      FUNCTION STDDVN(X,N)
      DIMENSION X(1)
      SUMX=0.
      SUMX2=0.
      DO 100 I=1,N
      SUMX=SUMX+X(I)
100   SUMX2=SUMX2+X(I)*X(I)
      STDDVN=SQRT((SUMX2-SUMX*SUMX/N)/(N-1))
      RETURN
      END
```

```
      REAL*8 GAMMA1(100,20),GAMMA2(100,20),GAMMA3(100,20),MATRIX(100,20)
     *,MEAN(100,2)/200*0.0/,TOBS( 100),A,X,FEW*4(100)/100*0.0/,ANS*4(4)/
     *4*0.0/,UBO*4(3)/3*0.0/
      COMMON CLEAN(SMIN,SMAX,XM,XL,ICT,ICTA,ICT1,ICT2,ICT3,ICT4,ICT5,ICT
     *6,ICTA1,ICTA2,ICTA3,ICTA4,ICTA5,ICTA6,S)
      INTEGER ACR(50),IPT(50,2)/100*0/,IM( 50)/ 50*0/,IBLK/2H  /,SYM/2H
     **/,ACRA(100),SYMA/1H*/,  INTVAL(8)/8*0/,S
      INTEGER GR1,GR2,GR3,GTGR3,GL1,GL2,GL3,GTGL3,MR
      DIMENSION STOR(100000), D(40), DA(40)
      B1=3.
      K=0
      K2=0
      MR=0
      K1=4
      IX=53693
C   DO LOOP FOR VARIOUS SAMPLE SIZES
C   THIS PROGRAMME UTILIZES FOUR SAMPLE SIZES: 5,10,15,20
      DO 778 IN=1,4
      K2=3
      MR=MR+5
      NN=0
C   DO LOOP FOR VARIOUS VALUES OF THE PARAMETERS IN EACH CYCLE
      DO 777 IB=1,5
      K2=K2+1
      MMR=MR*5000
      CD=(1.0*K1)/(K1+K2)
      DO 1000 II=1,50
      DO 907 IIA=1, 50
  907 IM(IIA)=0
C   GENERATION OF  FIRST ERLANG DISTRIBUTION
      DO 10 I=1,100
      DO 10 J=1,MR
      TR=0.0
      DO 11 IK=1,K1
      IY=IX*65539
      IF (IY) 5,6,6
    5 IY=IY+2147483647+1
    6 YFL=IY
      YFL=YFL*.4656613E-9
      IX=IY
   11 TR=TR+ALOG(YFL)
      X=(-1.0/B1)*(TR)
      IF (X.LT.XMAX) XMA=X
      IF (X.GT.XLA) XLA=X
   10 GAMMA1(I,J) = X
C GENERATION OF SECOND ERLANG DISTRIBUTION
      YFL = 0.0
      DO 13 I=1,100
```

```
      DO 13 J=1,MR
      TR=0.0
      DO 12 JJ = 1,K2
      IY =  IX*65539
      IF(IY) 7,8,8
7     IY=IY+2147483647+1
8     YFL=IY
      YFL = YFL*.4656613E-9
      IX = IY
   12 TR=TR+ALOG(YFL)
      X=(-1.0/B1)*(TR)
      IF(X.LT.XL) XL=X
      IF(X.GT.XM) XM=X
   13 GAMMA2(I,J) =X
C    THE GENERATION OF BETA DISTRIBUTION IS AS FOLLOWS
      DO 15 I=1,100
      DO 15 J= 1,MR
      GAMMA3(I,J) = GAMMA1(I,J)+ GAMMA2(I,J)
   15 MATRIX(I,J) = GAMMA1(I,J)/GAMMA3(I,J)
 9050 DO 905 I=1,100
      DO 905 J=1,MR
      S=S+1
      IF(MATRIX(I,J).LT.SMIN)SMIN=MATRIX(I,J)
      IF(MATRIX(I,J).GT.SMAX)SMAX=MATRIX(I,J)
  905 STOR(S)=MATRIX(I,J)
      DO 250 I=1,100
      DO 250 J=1,MR
      MEAN(I,1)=MEAN(I,1)+MATRIX(I,J)
  250 MEAN(I,2)=MEAN(I,2)+MATRIX(I,J)**2
      DO 87 I=1,100
      MEAN(I,2)=DSQRT((MEAN(I,2)-MEAN(I,1)*MEAN(I,1)/MR)/(MR-1))
   87 MEAN(I,1)=MEAN(I,1)/MR
      XM=MEAN(1,1)
      XL=MEAN(1,1)
C    COMPARISON OF THE SIGNIFICANCE VALUES IS AS FOLLOWS
      IF(MR.EQ.5) GO TO 26
      IF(MR.EQ.10) GO TO 27
      IF(MR.EQ.15) GO TO 28
      IF(MR.EQ.20) GO TO 29
   26 A=CD
      DO 20 I=1,100
      TOBS(I)=(MEAN(I,1)- A)/(MEAN(I,2)/MR**.5)
      IF(DABS(TOBS(I)).LE.0.271)ICT=ICT+1
      IF(DABS(TOBS(I)).GT.0.271) ICTA=ICTA+1
      IF(DABS(TOBS(I)).LE.0.741) ICT1=ICT1+1
      IF(DABS(TOBS(I)).GT.0.741) ICTA1=ICTA1+1
      IF(DABS(TOBS(I)).LE.1.533) ICT2=ICT2+1
      IF(DABS(TOBS(I)).GT.1.533) ICTA2=ICTA2+1
```

```
      IF(DABS(TOBS(I)).LE.2.132) ICT3=ICT3+1                            RIZ00920
      IF(DABS(TOBS(I)).GT.2.132)ICTA3=ICTA3+1                           RIZ00930
      IF(DABS(TOBS(I)).LE.2.776) ICT4=ICT4+1                            RIZ00940
      IF(DABS(TOBS(I)).GT.2.776)ICTA4=ICTA4+1                           RIZ00950
      IF(DABS(TOBS(I)).LE.3.747) ICT5=ICT5+1                            RIZ00960
      IF(DABS(TOBS(I)).GT.3.747)ICTA5=ICTA5+1                           RIZ00970
      IF(DABS(TOBS(I)).LE.4.604) ICT6=ICT6+1                            RIZ00980
      IF(DABS(TOBS(I)).GT.4.604)ICTA6=ICTA6+1                           RIZ00990
      IF(DABS(TOBS(I)).LE.8.610) ICT7=ICT7+1                            RIZ01000
      IF(DABS(TOBS(I)).GT.8.610)ICTA7=ICTA7+1                           RIZ01010
      IF(MEAN(I,1).LT.XM)XM=MEAN(I,1)                                   RIZ01020
      IF(MEAN(I,1).GT.XL)XL=MEAN(I,1)                                   RIZ01030
   20 CONTINUE                                                          RIZ01040
      GO TO 30                                                          RIZ01050
   27 A=CD                                                              RIZ01060
      DO 23 I=1,100                                                     RIZ01070
      TOBS(I)=(MEAN(I,1)- A)/(MEAN(I,2)/MR**.5)                         RIZ01080
      IF(DABS(TOBS(I)).LE.0.261)ICT=ICT+1                               RIZ01090
      IF(DABS(TOBS(I)).GT.0.261) ICTA=ICTA+1                            RIZ01100
      IF(DABS(TOBS(I)).LE.0.703) ICT1=ICT1+1                            RIZ01110
      IF(DABS(TOBS(I)).GT.0.703) ICTA1=ICTA1+1                          RIZ01120
      IF(DABS(TOBS(I)).LE.1.383) ICT2=ICT2+1                            RIZ01130
      IF(DABS(TOBS(I)).GT.1.383) ICTA2=ICTA2+1                          RIZ01140
      IF(DABS(TOBS(I)).LE.1.833) ICT3=ICT3+1                            RIZ01150
      IF(DABS(TOBS(I)).GT.1.833)ICTA3=ICTA3+1                           RIZ01160
      IF(DABS(TOBS(I)).LE.2.262) ICT4=ICT4+1                            RIZ01170
      IF(DABS(TOBS(I)).GT.2.262)ICTA4=ICTA4+1                           RIZ01180
      IF(DABS(TOBS(I)).LE.2.821) ICT5=ICT5+1                            RIZ01190
      IF(DABS(TOBS(I)).GT.2.821)ICTA5=ICTA5+1                           RIZ01200
      IF(DABS(TOBS(I)).LE.3.250) ICT6=ICT6+1                            RIZ01210
      IF(DABS(TOBS(I)).GT.3.250)ICTA6=ICTA6+1                           RIZ01220
      IF(DABS(TOBS(I)).LE.4.781) ICT7=ICT7+1                            RIZ01230
      IF(DABS(TOBS(I)).GT.4.781)ICTA7=ICTA7+1                           RIZ01240
   23 CONTINUE                                                          RIZ01250
      GO TO 30                                                          RIZ01260
   28 A=CD                                                              RIZ01270
      DO 24 I=1,100                                                     RIZ01280
      TOBS(I)=(MEAN(I,1)- A)/(MEAN(I,2)/MR**.5)                         RIZ01290
      IF(DABS(TOBS(I)).LE.0.258)ICT=ICT+1                               RIZ01300
      IF(DABS(TOBS(I)).GT.0.258) ICTA=ICTA+1                            RIZ01310
      IF(DABS(TOBS(I)).LE.0.692) ICT1=ICT1+1                            RIZ01320
      IF(DABS(TOBS(I)).GT.0.692) ICTA1=ICTA1+1                          RIZ01330
      IF(DABS(TOBS(I)).LE.1.345) ICT2=ICT2+1                            RIZ01340
      IF(DABS(TOBS(I)).GT.1.345) ICTA2=ICTA2+1                          RIZ01350
      IF(DABS(TOBS(I)).LE.1.761) ICT3=ICT3+1                            RIZ01360
      IF(DABS(TOBS(I)).GT.1.761)ICTA3=ICTA3+1                           RIZ01370
      IF(DABS(TOBS(I)).LE.2.145) ICT4=ICT4+1                            RIZ01380
      IF(DABS(TOBS(I)).GT.2.145)ICTA4=ICTA4+1                           RIZ01390
```

```
      IF(DABS(TOBS(I)).LE.2.624)  ICT5=ICT5+1
      IF(DABS(TOBS(I)).GT.2.264) ICTA5=ICTA5+1
      IF(DABS(TOBS(I)).LE.2.977)  ICT6=ICT6+1
      IF(DABS(TOBS(I)).GT.2.977) ICTA6=ICTA6+1
      IF(DABS(TOBS(I)).LE.4.140)  ICT7=ICT7+1
      IF(DABS(TOBS(I)).GT.4.140) ICTA7=ICTA7+1
 24   CONTINUE
      GO TO 30
 29   A=CD
      DO 25 I=1,100
      TOBS(I)=(MEAN(I,1)- A)/(MEAN(I,2)/MR**.5)
      IF(DABS(TOBS(I)).LE.0.257) ICT=ICT+1
      IF(DABS(TOBS(I)).GT.0.257)  ICTA=ICTA+1
      IF(DABS(TOBS(I)).LE.0.688)  ICT1=ICT1+1
      IF(DABS(TOBS(I)).GT.0.688)  ICTA1=ICTA1+1
      IF(DABS(TOBS(I)).LE.1.328)  ICT2=ICT2+1
      IF(DABS(TOBS(I)).GT.1.328)  ICTA2=ICTA2+1
      IF(DABS(TOBS(I)).LE.1.729)  ICT3=ICT3+1
      IF(DABS(TOBS(I)).GT.1.729) ICTA3=ICTA3+1
      IF(DABS(TOBS(I)).LE.2.093) ICT4=ICT4+1
      IF(DABS(TOBS(I)).GT.2.093) ICTA4=ICTA4+1
      IF(DABS(TOBS(I)).LE.2.539)  ICT5=ICT5+1
      IF(DABS(TOBS(I)).GT.2.539) ICTA5=ICTA5+1
      IF(DABS(TOBS(I)).LE.2.861)  ICT6=ICT6+1
      IF(DABS(TOBS(8)).GT.2.861) ICTA6=ICTA6+1
      IF(DABS(TOBS(I)).LE.3.883)  ICT7=ICT7+1
      IF(DABS(TOBS(I)).GT.3.883) ICTA7=ICTA7+1
 25   CONTINUE
 30   LOU=0.0
      DO 1000 I=1,100
      X=MEAN(I,2)**2
 901  MEAN(I,1)=0.0
1000  MEAN(I,2)=0.0
      A1=0.8
      A2=0.5
      A3=0.2
      A4=0.1
      A5=0.05
      A6=0.02
      A7=0.01
      A8=0.001
      Y1=ICTA/5000.
      Y2=ICTA1/5000.
      Y3=ICTA2/5000.
      Y4=ICTA3/5000.
      Y5=ICTA4/5000.
      Y6=ICTA5/5000.
      Y7=ICTA6/5000.
```

```
      Y8=ICTA7/5000.                                                    RIZ01880
      B2=5000                                                           RIZ01890
      F1=(ICT-.2*B2)**2/(.2*B2)                                         RIZ01900
      F2=(ICTA-A1*B2)**2/(A1*B2)                                        RIZ01910
      C1=F1+F2                                                          RIZ01920
      F3=(ICT1-.5*B2)**2/(.5*B2)                                        RIZ01930
      F4=(ICTA1-.5*B2)**2/(.5*B2)                                       RIZ01940
      C2=F3+F4                                                          RIZ01950
      F5=(ICT2-.8*B2)**2/(.8*B2)                                        RIZ01960
      F6=(ICTA2-.2*B2)**2/(.2*B2)                                       RIZ01970
      C3=F5+F6                                                          RIZ01980
      F7=(ICT3-.9*B2)**2/(.9*B2)                                        RIZ01990
      F8=(ICTA3-.1*B2)**2/(.1*B2)                                       RIZ02000
      C4=F7+F8                                                          RIZ02010
      F9=(ICT4-.95*B2)**2/(.95*B2)                                      RIZ02020
      F10=(ICTA4-.05*B2)**2/(.05*B2)                                    RIZ02030
      C5=F9+F10                                                         RIZ02040
      F11=(ICT5-.98*B2)**2/(.98*B2)                                     RIZ02050
      F12=(ICTA5-.02*B2)**2/(.02*B2)                                    RIZ02060
      C6=F11+F12                                                        RIZ02070
      F13=(ICT6-.99*B2)**2/(.99*B2)                                     RIZ02080
      F14=(ICTA6-.01*B2)**2/(.01*B2)                                    RIZ02090
      C7=F13+F14                                                        RIZ02100
      F15=(ICT7-.999*B2)**2/(.999*B2)                                   RIZ02110
      F16=(ICTA7-.001*B2)**2/(.001*B2)                                  RIZ02120
      C8=F15+F16                                                        RIZ02130
      DA(NN+IB)=C1                                                      RIZ02140
      DA(NN+IB+1)=C2                                                    RIZ02150
      DA(NN+IB+2)=C3                                                    RIZ02160
      DA(NN+IB+3)=C4                                                    RIZ02170
      DA(NN+IB+4)=C5                                                    RIZ02180
      DA(NN+IB+5)=C6                                                    RIZ02190
      DA(NN+IB+6)=C7                                                    RIZ02200
      DA(NN+IB+7)=C8                                                    RIZ02210
      D(NN+IB)=Y1                                                       RIZ02220
      D(NN+IB+1)=Y2                                                     RIZ02230
      D(NN+IB+2)=Y3                                                     RIZ02240
      D(NN+IB+3)=Y4                                                     RIZ02250
      D(NN+IB+4)=Y5                                                     RIZ02260
      D(NN+IB+5)=Y6                                                     RIZ02270
      D(NN+IB+6)=Y7                                                     RIZ02280
      D(NN+IB+7)=Y8                                                     RIZ02290
      NN=NN+7                                                           RIZ02300
      CQ1=.098                                                          RIZ02310
      CQ2=.455                                                          RIZ02320
      CQ3=1.324                                                         RIZ02330
      CQ4=2.71                                                          RIZ02340
      CQ5=3.84                                                          RIZ02350
```

```
      CQ6=5.02
      CQ7=6.63
      CQ8=7.88
C*************************************************************************
      XBAR1=XBAR(STOR,MMR)
      SIGMA=STDDVN(STOR,MMR)
      BASE=-3.0
      DO 111 I=1,8
  111 INTVAL(I)=0
      DO 21 I=1,MMR
      IF((STOR(I)-XBAR1)/SIGMA .LE. BASE) INTVAL(1)=INTVAL(1)+1
      DO 22 J=2,7
      K=J-2
      IF((STOR(I)-XBAR1)/SIGMA .GT.BASE+K .AND.(STOR(I)-XBAR1)/SIGMA .LE
     *. BASE+(K+1)) INTVAL(J)=INTVAL(J)+1
   22 CONTINUE
      IF(( STOR(I)-XBAR1)/SIGMA .GT. BASE+6) INTVAL(8)=INTVAL(8)+1
   21 CONTINUE
C IF THE DISTRIBUTION WAS NORMAL THEN THE FOLLOWING GROUPINGS WOULD BE
CEXLECTED
      N=MMR
      GR1=(0.34)*N
      GR2=(0.136)*N
      GR3=(0.021 )*N
      GTGR3=(0.003 )*N
      GL1=(0.34)*N
      GL2=(0.136)*N
      GL3=(0.021 )*N
      GTGL3=(0.003 )*N
C*****************************************************Q*******************
C*****************************************************
C     THE VARIOUS WRITE STATEMENTS ARE AS FOLLOWS
      WRITE(6,41) GTGL3,GL3,GL2,GL1,GR1,GR2,GR3,GTGR3
      WRITE(6,38) C1,C2,C3,C4,C5,C6,C7,C8
      WRITE(6, 41)INTVAL
      WRITE(6,62)
      WRITE(6,130)
      WRITE(6,131)
      WRITE(6,601)
      WRITE(6,602) MMR
      WRITE(6,603) Y1,Y2,Y3,Y4,Y5,Y6,Y7,Y8
      WRITE(6,603) A1,A2,A3,A4,A5,A6,A7,A8
      WRITE(6,604) A
      WRITE(6,771) SIGMA
      WRITE(6,772) XBAR1
      WRITE(6,779)MR
   52 FORMAT('0',10X,10F12.8)
  800 FORMAT('0','  BETA RANDOM NUMBERS ARE AS FOLLOWS ')
```

```
 100 FORMAT('0',' GAMM2 RANDOM NUMBERS ARE AS FOLLOWS ')
 99  FORMAT('0',' GAMM1 RANDOM NUMBERS ARE AS FOLLOWS ')
 101 FORMAT('0', ' IX VALUE = ',I15,' YFL VALUE = ',F12.8)
 130 FORMAT(1H1,' THE BETA FREQUENCY INTERVALS ARE AS FOLLOWS ',//)
 131 FORMAT(10X,'  LE M-3S ,  LE M-2S ,  LE M-S ,  LE MEW ,  LE M+S ,
    *LE M+2S  ,  LE M+3> ,  GT M+3S  ',//)
 41  FORMAT(10X,8I8)
 60  FORMAT(1HO,' NUMBER OF TOBS LE TCRIT AT ALPHA =0.8 ',I5, ' NUMBER
    *OF TOBS GT TCRIT ',I5,//)
 31  FORMAT(1HO,' NUMBER OF TOBS LE TCRIT AT ALPHA = 0.5',I5, ' NUMBER
    *OF TOBS GT TCRIT ',I5,//)
 32  FORMAT(1HO,' NUMBER OF TOBS LE TCRIT AT ALPHA =0.2 ',I5, ' NUMBER
    *OF TOBS GT TCRIT ',I5,//)
 33  FORMAT(1HO,' NUMBER OF TOBS LE TCRIT AT ALPHA =.1  ',I5, ' NUMBER
    *OF TOBS GT TCRIT ',I5,//)
 34  FORMAT(1HO,' NUMBER OF TOBS LE TCRIT AT ALPHA =.05 ',I5, ' NUMBER
    *OF TOBS GT TCRIT ',I5,//)
 35  FORMAT(1HO,' NUMBER OF TOBS LE TCRIT AT ALPHA =0.02',I5, ' NUMBER
    *OF TOBS GT TCRIT ',I5,//)
 36  FORMAT(1HO,' NUMBER OF TOBS LE TCRIT AT ALPHA =0.01',I5, ' NUMBER
    *OF TOBS GT TCRIT ',I5,//)
 37  FORMAT(1HO,' NUMBER OF TOBS LE TCRIT AT ALPHA = 0.001',I5,'NUMBER
    *GT TCRIT ',I5,//)
 62  FORMAT('0',' IF THE DISTRIBUTION WAS NORMAL THE FOLLOWING WOULD BE
    *EXPECTED ',//)
 601 FORMAT(1HO,'  DISTRIBUTION: BETA ',//)
 779 FORMAT(1HO,' SAMPLE SIZE = ',I5,//)
 604 FORMAT(1HO,' THEORETICAL SAMPLE MEAN = ',F5.2,//)
 771 FORMAT(1HO,'STANDARD DEVIATION = ',F10.5,//)
 772 FORMAT(1HO,' D MEAN= ',F10.5,//)
 602 FORMAT(1HO,' TOTAL DATA POINTS = ',I10,//)
 603 FORMAT(10X,8F12.3,//)
 38  FORMAT(10X,8F13.3)
C*************************************************************************
     DO 7009 I=1,100
7009 TOBS(I)=0
 777 CONTINUE
     WRITE(6,609)
 609 FORMAT(1H1,' COMPARISON OF SIGNIFICANCE LEVELS IS AS FOLLOWS',//)
     WRITE(6,381)
 381 FORMAT(1HO,' CRITICAL VALUES ARE ',//)
     WRITE(6,603) A1,A2,A3,A4,A5,A6,A7,A8
     WRITE(6,382)
 382 FORMAT(1HO,' EMPIRICAL VALUES ARE ',//)
     WRITE(6,603) D
     WRITE(6.155)
 155 FORMAT(1HO,'CHISQUARE CRITICAL VALUES ARE',//)
     WRITE(6,603) DA
```

```
      WRITE(6,603)CQ1,CQ2,CQ3,CQ4,CQ5,CQ6,CQ7,CQ8                       RI703320
      WRITE(6,156)                                                      RI703330
  156 FORMAT(1H0,'CHI SQ STATISTICS IS AS FOLLOWS',//)                  RI703340
C**************************************************0********************RI703350
 778  CONTINUE                                                          RI703360
      WRITE(6,7001)                                                     RI703370
 7001 FORMAT(1H1,20X,'END OF JOB')                                      RI703380
      STOP                                                              RI703390
      END                                                               RI703400
C**************************************************0********************RI703410


      SUBROUTINE CLEAN                                                  RI703420
      COMMON CLEAN(SMIN,SMAX,XM,XL,ICT,ICTA,ICT1,ICT2,ICT3,ICT4,ICT5,ICTRI703430
     *6,ICTA1,ICTA2,ICTA3,ICTA4,ICTA5,ICTA6,S)                          RI703440
      SMIN=0                                                            RI703450
      SMAX=0                                                            RI703460
      XM=0                                                              RI703470
      XL=0                                                              RI703480
      ICT=0                                                             RI703490
      ICT1=0                                                            RI703500
      ICT2=0                                                            RI703510
      ICT3=0                                                            RI703520
      ICT4=0                                                            RI703530
      ICT5=0                                                            RI703540
      ICT6=0                                                            RI703550
      ICT7=0                                                            RI703560
      ICTA=0                                                            RI703570
      ICTA1=0                                                           RI703580
      ICTA2=0                                                           RI703590
      ICTA3=0                                                           RI703600
      ICTA4=0                                                           RI703610
      ICTA5=0                                                           RI703620
      ICTA6=C                                                           RI703630
      ICTA7=0                                                           RI703640
      S=0                                                               RI703650
      RETURN                                                            RI703660
      END                                                               RI703670
C*********************************************************************  RI703680


      FUNCTION XBAR(X,N)                                                RI703690
      DIMENSION X(1)                                                    RI703700
      SUMX=0.                                                           RI703710
      DO 100 I=1,N                                                      RI703720
100   SUMX=SUMX+X(I)                                                    RI703730
      XBAR=SUMX/N                                                       RI703740
      R5TURN                                                            RI703750
```

```
      END                                                                RIZ03760
C*********************************************************************** RIZ03770

      FUNCTION STDDVN(X,N)                                               RIZ03780
      DIMENSION X(1)                                                     RIZ03790
      SUMX=0.                                                            RIZ03800
      SUMX2=0.                                                           RIZ03810
      DO 100 I=1,N                                                       RIZ03820
      SUMX=SUMX+X(I)                                                     RIZ03830
100   SUMX?=SUMX2+X(I)*X(I)                                              RIZ03840
      STDDVN=SQRT((SUMX2-SUMX*SUMX/N)/(N-1))                             RIZ03850
      RETURN                                                             RIZ03860
      END                                                                RIZ03870
C*****************************************************0*****************RIZ03880
```

APPENDIX J

Simulation 1, Cycle 1

The beta frequency intervals are as follows

LE M-3S, LE M-2S, LE M-S, LE MEW, LE M+S, LE M+2S, LE M+3>, GT M+3S

0 411 3879 8062 8333 3896 419 0 (1)

If the distribution was normal the following would be expected

74 524 3399 8499 8499 3399 524 74 (2)

Distribution: Beta

Sample Size = 5 (3)

Theoretical Sample Mean = 0.50 (4)

Standard Deviation = 0.16731

D Mean = 0.49941

Total Data Points = 25000

Actual α = 0.800 0.500 0.200 0.100 0.050 0.020 0.010 0.001 (5)

Observed α= 0.793 0.485 0.196 0.099 0.051 0.020 0.011 0.002 (6)

$\sum_{i=1}^{2} \frac{(o_i-e_i)^2}{e_i}$ = 1.531 4.743 0.405 0.020 0.152 0.000 0.182 1.802 (7)

Simulation 1, Cycle 2

The Beta Frequency Intervals are as follows

LE M-3S,	LE M-2S,	LE M-S,	LE MEW,	LE M+S,	LE M+2S,	LE M+3>,	GT M+3S	
0	298	3894	8414	8146	3670	575	0	(1)

If the distribution was normal the following would be expected

74	524	3399	8499	8499	3399	524	74	(2)

Distribution: Beta

Sample Size = 5 (3)

Theoretical Sample Mean = 0.50 (4)

Standard Deviation = 0.16731

D Mean = 0.44258

Total Data Points = 25000

Actual α =	0.800	0.500	0.200	0.100	0.050	0.020	0.010	0.001	(5)
Observed α=	0.793	0.480	0.190	0.099	0.050	0.021	0.011	0.001	(6)
$\sum_{i=1}^{2} \frac{(o_i - e_i)^2}{e_i}$ =	1.361	8.000	3.125	0.056	0.017	0.092	0.182	0.000	(7)

Simulation 1, Cycle 3

The Beta frequency intervals are as follows

LE M-3S, LE M-2S, LE M-S, LE MEW, LE M+S, LE M+2S, LE M+3>, GT M+3S

0 194 3890 8739 7987 3515 660 15 (1)

If the distribution was normal the following would be expected

74 524 3399 8499 8499 3399 524 74 (2)

Distribution: Beta

Sample Size = 5 (3)

Theoretical Sample Mean = 0.40 (4)

Standard Deviation = 0.15020

D Mean = 0.39854

Total Data Points = 25000

Actual α =	0.800	0.500	0.200	0.100	0.050	0.020	0.010	0.001	(5)
Observed α=	0.792	0.488	0.192	0.096	0.050	0.020	0.010	0.001	(6)
$\sum_{i=1}^{2} \frac{(o_i-e_i)^2}{e_i}$ =	2.000	2.977	2.000	0.720	0.017	0.041	0.000	0.000	(7)

Simulation 1, Cycle 4

The Beta frequency intervals are as follows

LE M-3S, LE M-2S, LE M-S, LE MEW, LE M+S, LE M+2S, LE M+3>, GT M+3S

0 151 3942 8919 7849 3445 656 38 (1)

If the distribution was normal the following would be expected

74 524 3399 8499 8499 3399 524 74 (2)

Distribution: Beta

Sample Size = 5 (3)

Theoretical Sample Mean = 0.36 (4)

Standard Deviation = 0.14029

D Mean = 0.36259

Total Data Points = 25000

Actual α =	0.800	0.500	0.200	0.100	0.050	0.020	0.010	0.001	(5)
Observed α=	0.798	0.492	0.204	0.098	0.053	0.020	0.010	0.001	(6)
$\sum_{i=1}^{2} \frac{(o_i - e_i)^2}{e_i}$ =	0.080	1.217	0.405	0.269	0.825	0.000	0.000	0.000	(7)

Simulation 1, Cycle 5

The Beta frequency intervals are as follows

LE M-3S, LE M-2S, LE M-S, LE MEW, LE M+S, LE M+2S, LE M+3>, GT M+3S

0 100 3919 9111 7736 3382 696 56 (1)

If the distribution was normal the following would be expected

74 524 3399 8499 8499 3399 524 74 (2)

Distribution: Beta

Sample Size = 5 (3)

Theoretical Sample Mean = 0.33 (4)

Standard Deviation = 0.13295

D Mean = 0.33244

Total Data Points = 25000

Actual α = 0.800 0.500 0.200 0.100 0.050 0.020 0.010 0.001 (5)

Observed α= 0.801 0.504 0.205 0.107 0.061 0.023 0.012 0.003 (6)

$\sum_{i=1}^{2} \frac{(o_i-e_i)^2}{e_i}$ = 0.061 0.353 0.781 2.569 12.737 2.612 2.020 12.813 (7)

Consolidate Results of Simulation 1

Comparison of significance levels is as follows

Actual Alpha values are

0.800	0.500	0.200	0.100	0.050	0.020	0.010	0.001	5	

Empirical Alpha values are

0.793	0.485	0.196	0.099	0.051	0.020	0.011	0.002	6	Cycle 1
0.793	0.480	0.190	0.099	0.050	0.021	0.011	0.001	6	Cycle 2
0.792	0.488	0.192	0.096	0.050	0.020	0.010	0.001	6	Cycle 3
0.798	0.492	0.204	0.098	0.053	0.020	0.010	0.001	6	Cycle 4
0.801	0.504	0.205	0.107	0.061	0.023	0.012	0.003	6	Cycle 5

Chisquare Critical values are

$\chi^2_{1,.95}$

0.098	0.455	1.324	2.710	3.840	5.020	6.630	7.880

Chi Sq Statistics is as follows

1.531	4.743	0.405	0.020	0.152	0.000	0.182	1.802	7	Cycle 1
1.361	8.000	3.125	0.056	0.017	0.092	0.182	0.000	7	Cycle 2
2.000	2.977	2.000	0.720	0.017	0.041	0.000	0.000	7	Cycle 3
0.080	1.217	0.405	0.269	0.825	0.000	0.000	0.000	7	Cycle 4
0.061	0.353	0.781	2.569	12.737	2.612	2.020	12.813	7	Cycle 5

BIBLIOGRAPHY

1. Gosset, W.S., "The Probable Error of the Mean," Biometrika, v.6, p. 1-25.

2. Fisher, R.A., Statistical Methods for Research Workers, Hafner Book Company, 1958.

3. Fisher, R.A., "Probable Error," Metron, v.1, p. 1-32.

4. Pearson, E.S., "Effects of Excess and Skewness," Biometrika, v.20, p. 175-209.

5. Gayden, A.K., "The Distribution of Student's T in Random Samples From Non-Normal Universes," Biometrika, v.36, p. 111-121.

6. Pearson, E.S., "Various Properties of Student's T," Biometrika, v.21, p. 121-162.

7. Box, G.E.P., "Non-Normality and Test on Variance," Journal of the Royal Statistical Society, p. 318-319, 1953.

8. Geary, R.C., "The Distribution of Student's Ratio for Non-Normal Samples," Journal of the Royal Statistical Society, v.3, p. 178-180, 1936.

9. Cucconio, O., "Critical Values of 'Student' t, Annals of Mathematical Statistics, v.5, p. 1121-1125, 1964.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
3.	Department of Operations Analysis, Code 55 Naval Postgraduate School Monterey, California 93940	1
4.	Asst. Professor G. A. Tuck, Code 55Tk Department of Operations Analysis Naval Postgraduate School Monterey, California 93940	1
5.	MAJOR M. R. Nomani, USA 1222 Spruance Monterey, California 93940	2

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Naval Postgraduate School Monterey, California 93940	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

An Investigation of the Robustness of "Student's" t-Test Using Beta and Erlang Distributions

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Master's Thesis, September 1970

5. AUTHOR(S) (First name, middle initial, last name)

M. Rizwan Nomani

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
September 1970	88	9

8a. CONTRACT OR GRANT NO. b. PROJECT NO. c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Naval Postgraduate School Monterey, California 93940

13. ABSTRACT

The robustness of "Student's" t-test with regard to the assumption of normality is investigated. This is accomplished by empirically developing the distribution of a new statistic from a large number of samples of varying sizes from the Beta and the gamma distributions using different values of the parameters. The various significance levels of this new distribution are then compared with the corresponding significance levels from "Student's" t distribution using an IBM 360 computer. A comparison of the distribution frequencies with the standard normal distribution frequencies is also presented.

KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Robustness of a test						
Power of the test						
"Student's" t-test						
Mode and variance						
Critical path scheduling						
Inverse probability integral transformation						
Simulation techniques						
Critical values						
Frequency comparison						
Leptokurtic curve						
Symmetric distribution						
Skewed distribution						

119307
Thesis
N845
c.1
Nomani
An investigation of the robustness of "Student's" t-test using Beta and Erlang distributions.